GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 10:03:28 ; Search time 16.5 Seconds
(without alignments)
137.669 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	44	3	US-08-955-636-24	Sequence 24, Appl
2	179	90.9	419	4	US-10-182-263-6	Sequence 6, Appli
3	176	89.3	44	3	US-08-955-636-35	Sequence 35, Appl
4	174	88.3	419	4	US-10-182-263-3	Sequence 3, Appli
5	174	88.3	419	4	US-10-182-263-4	Sequence 4, Appli
6	174	88.3	419	4	US-10-182-263-5	Sequence 5, Appli
7	173	87.8	44	3	US-08-955-636-20	Sequence 20, Appl
8	170	86.3	44	3	US-08-955-636-21	Sequence 21, Appl
9	168	85.3	44	3	US-08-955-636-19	Sequence 19, Appl
10	168	85.3	44	3	US-08-955-636-22	Sequence 22, Appl
11	160	81.2	44	3	US-08-955-636-1	Sequence 1, Appli
12	160	81.2	44	3	US-08-955-636-25	Sequence 25, Appl
13	160	81.2	45	2	US-08-965-832-2	Sequence 2, Appli
14	160	81.2	419	1	US-08-295-411-1	Sequence 1, Appli
15	160	81.2	419	2	US-08-955-471-1	Sequence 1, Appli
16	160	81.2	419	4	US-09-667-570A-3	Sequence 3, Appli
17	160	81.2	419	4	US-10-182-263-1	Sequence 1, Appli
18	160	81.2	419	5	PCT-US92-10242-1	Sequence 1, Appli
19	160	81.2	460	2	US-08-756-506-2	Sequence 2, Appli
20	160	81.2	460	2	US-08-756-506-4	Sequence 4, Appli
21	160	81.2	460	6	5270178-13	Patent No. 5270178
22	160	81.2	460	6	5270178-14	Patent No. 5270178
23	160	81.2	460	6	5270178-15	Patent No. 5270178
24	160	81.2	460	6	5270178-16	Patent No. 5270178
25	160	81.2	461	4	US-10-182-263-2	Sequence 2, Appli
26	160	81.2	461	6	5225537-2	Patent No. 5225537
27	160	81.2	461	6	5270178-17	Patent No. 5270178
28	160	81.2	461	6	5270178-18	Patent No. 5270178
29	160	81.2	461	6	5460953-3	Patent No. 5460953
30	147	74.6	42	2	US-08-745-254A-2	Sequence 2, Appli
31	147	74.6	461	6	5270178-2	Patent No. 5270178
32	143	72.6	41	1	US-08-229-280-5	Sequence 5, Appli
33	138	70.1	42	4	US-09-383-667-8	Sequence 8, Appli
34	129	65.5	409	3	US-09-065-872-2	Sequence 2, Appli
35	129	65.5	409	4	US-09-667-570A-2	Sequence 2, Appli
36	129	65.5	410	3	US-09-065-872-1	Sequence 1, Appli
37	129	65.5	410	4	US-09-667-570A-1	Sequence 1, Appli
38	117	59.4	44	3	US-08-955-636-23	Sequence 23, Appl
39	116	58.9	44	3	US-08-955-636-2	Sequence 2, Appli
40	114	57.9	139	1	US-08-330-978-2	Sequence 2, Appli
41	114	57.9	139	1	US-08-474-042-2	Sequence 2, Appli
42	114	57.9	139	1	US-08-484-558-2	Sequence 2, Appli
43	114	57.9	139	1	US-08-774-592-2	Sequence 2, Appli
44	114	57.9	437	1	US-08-487-037-2	Sequence 2, Appli
45	114	57.9	437	1	US-08-487-037-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-955-636-24
; Sequence 24, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-24

Query Match 90.9%; Score 179; DB 3; Length 44;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||||||||||||||||||||||||||||||||||||||||||
Db 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44

RESULT 2
US-10-182-263-6
; Sequence 6, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-6

Query Match 90.9%; Score 179; DB 4; Length 419;
Best Local Similarity 79.5%; Pred. No. 2.8e-22;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 3
US-08-955-636-35
; Sequence 35, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-35

Query Match 89.3%; Score 176; DB 3; Length 44;
Best Local Similarity 97.7%; Pred. No. 6.9e-23;
Matches 43; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLXXLREGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44

RESULT 4
US-10-182-263-3
; Sequence 3, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-3

Query Match 88.3%; Score 174; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 5
US-10-182-263-4
; Sequence 4, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-4

Query Match 88.3%; Score 174; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 6
US-10-182-263-5
; Sequence 5, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-5

Query Match 88.3%; Score 174; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

```
RESULT 7
US-08-955-636-20
; Sequence 20, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-20

  Query Match             87.8%;  Score 173;  DB 3;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 2.2e-22;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||| ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 8
US-08-955-636-21
; Sequence 21, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-21

  Query Match             86.3%;  Score 170;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 7.2e-22;
  Matches   42;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||: ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 9
US-08-955-636-19
; Sequence 19, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-19

  Query Match             85.3%;  Score 168;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 1.6e-21;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 10
US-08-955-636-22
; Sequence 22, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-22

  Query Match             85.3%;  Score 168;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 1.6e-21;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRDSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 11
US-08-955-636-1
; Sequence 1, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-1

  Query Match             81.2%;  Score 160;  DB 3;  Length 44;
  Best Local Similarity   90.9%;  Pred. No. 3.6e-20;
  Matches   40;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||::||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 12
US-08-955-636-25
; Sequence 25, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-25

  Query Match             81.2%;  Score 160;  DB 3;  Length 44;
  Best Local Similarity   93.2%;  Pred. No. 3.6e-20;
  Matches   41;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  |||||||||||||||| ||||||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAFXIFEDVDDTLAFWSKH 44


RESULT 13
US-08-965-832-2
; Sequence 2, Application US/08965832
; Patent No. 5847085
;  GENERAL INFORMATION:
;    APPLICANT:  CHARLES T. ESMON AND MIKHAIL D. SMIRNOV
;    TITLE OF INVENTION:  Modified Protein C
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center, 1201 West
;      STREET:  Peachtree Street
;      CITY:  Atlanta
;      STATE:  GA
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/965,832
;      FILING DATE:  7-NOV-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/745,254
;      FILING DATE:  8-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/053,768
;      FILING DATE:  25-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF 165/167
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-873-8794
;      TELEFAX:  (404)-873-8795
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:
;      LOCATION:  6, 7, 14, 16, 19, 20, 25, 26, 29
;      OTHER INFORMATION:  /note= "where Xaa means gamma
;      OTHER INFORMATION:  carboxylglutamic acid"
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      OTHER INFORMATION:  /note= "partial sequence of human protein C"
US-08-965-832-2

  Query Match             81.2%;  Score 160;  DB 2;  Length 45;
  Best Local Similarity   90.9%;  Pred. No. 3.7e-20;
  Matches   40;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||::||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 14
US-08-295-411-1
; Sequence 1, Application US/08295411
; Patent No. 5679639
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5679639th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,411
;      FILING DATE:  22-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/793,989
;      FILING DATE:  18-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
```

```
;       REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-554-2937
;       TELEFAX:  619-554-6312
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  419 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1..157
;       OTHER INFORMATION:  /note= "Protein C Light Chain"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  158..169
;       OTHER INFORMATION:  /note= "Protein C Activation
;       OTHER INFORMATION:  Peptide"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  170..419
;       OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-295-411-1

  Query Match             81.2%;  Score 160;  DB 1;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 4.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  || | ||  ||||  || ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 15
US-08-955-471-1
; Sequence 1, Application US/08955471
; Patent No. 5968751
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5968751th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,471
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/295,411
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  419 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..157
;      OTHER INFORMATION:  /note= "Protein C Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  158..169
;      OTHER INFORMATION:  /note= "Protein C Activation
;      OTHER INFORMATION:  Peptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  170..419
;      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-955-471-1

  Query Match             81.2%;  Score 160;  DB 2;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 4.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  || | ||  ||||  || ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


Search completed: March  1, 2004, 10:12:21
Job time : 17.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:55:12 ; Search time 37.5 Seconds
(without alignments)
370.208 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	151	76.6	456	6	Q9TTR0	Q9ttr0 canis famil
2	140	71.1	460	11	Q91WN8	Q91wn8 mus musculu
3	134	68.0	460	11	Q99PC6	Q99pc6 mus musculu
4	118	59.9	55	4	Q8J002	Q8j002 homo sapien
5	118	59.9	55	4	Q8IXB5	Q8ixb5 homo sapien
6	115	58.4	482	11	Q63207	Q63207 rattus norv
7	109	55.3	455	13	Q7SY86	Q7sy86 xenopus lae
8	107	54.3	231	4	Q8N2N6	Q8n2n6 homo sapien
9	107	54.3	433	13	Q804X5	Q804x5 gallus gall
10	105	53.3	524	13	Q7SXH8	Q7sxh8 brachydanio
11	103.5	52.5	443	13	Q8JHC9	Q8jhc9 brachydanio
12	101	51.3	340	11	Q80Y26	Q80y26 mus musculu
13	101	51.3	481	11	O54740	O54740 mus musculu
14	101	51.3	481	11	Q99L32	Q99132 mus musculu
15	101	51.3	481	11	O88947	O88947 mus musculu
16	99.5	50.5	442	13	Q804X1	Q804x1 fugu rubrip
17	99	50.3	679	4	Q96PQ8	Q96pq8 homo sapien
18	97	49.2	474	13	Q8JHC8	Q8jhc8 brachydanio
19	95	48.2	469	6	Q9GMD9	Q9gmd9 ornithorhyn
20	93	47.2	229	13	Q8JJ40	Q8jj40 xenopus lae
21	93	47.2	434	13	Q7T3B6	Q7t3b6 brachydanio
22	92	46.7	268	4	Q8NEK6	Q8nek6 homo sapien
23	87	44.2	425	13	Q804X7	Q804x7 gallus gall
24	87	44.2	612	13	Q804W7	Q804w7 fugu rubrip
25	85	43.1	475	13	Q804W9	Q804w9 fugu rubrip
26	85	43.1	497	4	Q7Z715	Q7z715 homo sapien
27	85	43.1	650	4	Q16519	Q16519 homo sapien
28	85	43.1	650	4	Q9NSD0	Q9nsd0 homo sapien
29	84	42.6	100	4	Q15253	Q15253 homo sapien
30	84	42.6	622	4	Q7Z7P3	Q7z7p3 homo sapien
31	82.5	41.9	542	5	Q8T6I3	Q8t6i3 halocynthia
32	81	41.1	471	13	Q804X6	Q804x6 gallus gall
33	80	40.6	446	11	Q61109	Q61109 mus musculu
34	80	40.6	461	6	Q95ND7	Q95nd7 pan troglod
35	80	40.6	461	6	Q95ND6	Q95nd6 pan troglod
36	79	40.1	441	13	Q804X2	Q804x2 fugu rubrip
37	78	39.6	138	6	Q28994	Q28994 sus scrofa
38	78	39.6	607	13	Q91001	Q91001 gallus gall
39	78	39.6	648	6	Q29094	Q29094 sus scrofa
40	77	39.1	52	4	Q8IXD5	Q8ixd5 homo sapien
41	76	38.6	446	11	Q8K3U6	Q8k3u6 rattus norv
42	74	37.6	52	4	Q8IXC5	Q8ixc5 homo sapien
43	73.5	37.3	433	13	Q8JHD0	Q8jhd0 brachydanio
44	73.5	37.3	433	13	Q90YK1	Q90yk1 brachydanio
45	73	37.1	49	6	Q95ME8	Q95me8 bos taurus

ALIGNMENTS

RESULT 1
Q9TTR0
ID Q9TTR0 PRELIMINARY; PRT; 456 AA.
AC Q9TTR0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein C precursor.
GN PROC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Leeb T., Kopp T., Deppe A., Breen M., Matis U., Brunnberg L.,
RA Brenig B.;
RT "Molecular characterization and chromosomal assignment of the canine
RT protein C gene.";
RL Mamm. Genome 10:135-139(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99371952; PubMed=10443005;
RA Leeb T., Pfeiffer I., Kopp T., Deppe A., Brenig B.;
RT "Analysis of canine protein C gene polymorphisms.";
RL Anim. Genet. 30:237-238(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AJ001979; CAA05126.1; -.
DR HSSP; P04070; 1AUT.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.

```
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease; Signal.
FT   SIGNAL        1     42       POTENTIAL.
FT   CHAIN         43    192      PROTEIN C LIGHT CHAIN.
FT   CHAIN        193    194      PROTEIN C CONNECTING DIPEPTIDE.
FT   CHAIN        195    456      PROTEIN C HEAVY CHAIN.
SQ   SEQUENCE   456 AA;  50813 MW;  7AD3A8C1C34E59FF CRC64;

  Query Match             76.6%;  Score 151;  DB 6;  Length 456;
  Best Local Similarity   63.6%;  Pred. No. 3.8e-17;
  Matches   28;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  || ||::|||||:|||:
Db         43 ANSFLEEIRAGSLERECMEEICDFEEAKEIFQNVDDTLAYWSKY 86


RESULT 2
Q91WN8
ID   Q91WN8      PRELIMINARY;      PRT;   460 AA.
AC   Q91WN8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; BC013896; AAH13896.1; -.
DR   HSSP; P00761; 1AN1.
DR   MGD; MGI:97771; Proc.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   460 AA;  51818 MW;  0117F26E68FCC274 CRC64;

  Query Match             71.1%;  Score 140;  DB 11;  Length 460;
  Best Local Similarity   59.1%;  Pred. No. 2.9e-15;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  |: ||::|:||||||  |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
Q99PC6
ID   Q99PC6      PRELIMINARY;      PRT;   460 AA.
AC   Q99PC6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Anticoagulant protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL;
RA   Korf I.;
RT   "Complete sequence of UC72A01.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF318182; AAK07918.1; -.
DR   HSSP; P04070; 1AUT.
DR   MGD; MGI:97771; Proc.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
```

```
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   460 AA;  51784 MW;  0293BC25E9D3ED16 CRC64;

  Query Match             68.0%;  Score 134;  DB 11;  Length 460;
  Best Local Similarity   56.8%;  Pred. No. 3.1e-14;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  |||   |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDLEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
Q8J002
ID   Q8J002      PRELIMINARY;      PRT;    55 AA.
AC   Q8J002;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein C (Fragment).
GN   PROC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kinoshita S., Iida H., Inoue S., Watanabe K., Kurihara M., Wada Y.,
RA   Ono M., Dongchon K., Hamasaki N.;
RT   "Gene Analysis of Anticoagulation Factors in Japanese Thrombotic
RT   Patients. Genetic Background of Thrombophilia in Japan.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB083700; BAC21172.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
FT   NON_TER       1       1
FT   NON_TER      55      55
SQ   SEQUENCE   55 AA;  6527 MW;  4F89496534A78836 CRC64;

  Query Match             59.9%;  Score 118;  DB 4;  Length 55;
  Best Local Similarity   64.9%;  Pred. No. 1.7e-12;
  Matches   24;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDT 37
              |||||  ||  || | ||  |||| || ||::||||
Db         19 ANSFLKELRHSSLERECIEEICDFEEAKEIFQNVDDT 55


RESULT 5
Q8IXB5
ID   Q8IXB5      PRELIMINARY;      PRT;    55 AA.
AC   Q8IXB5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein C (Fragment).
GN   PROC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hamasaki S., Kang D., Kinoshita S., Iida K., Inoue S., Watanabe K.,
RA   Kurihara M., Wada Y., Ono M.;
RT   "Gene analysis of anticoagulation factors in Japanese thrombotic
RT   patients.Genetic background of thrombophilia in Japan.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB086851; BAC53631.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
FT   NON_TER       1       1
FT   NON_TER      55      55
SQ   SEQUENCE   55 AA;  6475 MW;  3803696534BC9289 CRC64;

  Query Match             59.9%;  Score 118;  DB 4;  Length 55;
  Best Local Similarity   64.9%;  Pred. No. 1.7e-12;
  Matches   24;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDT 37
              |||||  ||  || | ||  |||| || ||::||||
Db         19 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDT 55


RESULT 6
Q63207
ID   Q63207      PRELIMINARY;      PRT;   482 AA.
AC   Q63207;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Factor X.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=96093366; PubMed=8578539;
RA   Stanton C., Ross R.P., Hutson S., Wallin R.;
RT   "Evidence for competition between vitamin K-dependent clotting factors
RT   for intracellular processing by the vitamin K-dependent gamma-
RT   carboxylase.";
RL   Thromb. Res. 80:63-73(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; X79807; CAA56202.1; -.
DR   PIR; S49075; EXRT.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
```

```
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   482 AA;  54265 MW;  0284678E3954A698 CRC64;

  Query Match             58.4%;  Score 115;  DB 11;  Length 482;
  Best Local Similarity   43.2%;  Pred. No. 6e-11;
  Matches   19;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||   :::|:| | |:  || |  |: :||| : |  ||:|:
Db         41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 7
Q7SY86
ID   Q7SY86      PRELIMINARY;      PRT;   455 AA.
AC   Q7SY86;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RA   Klein S., Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054968; AAH54968.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   455 AA;  51811 MW;  07C027ED2B495330 CRC64;

  Query Match             55.3%;  Score 109;  DB 13;  Length 455;
  Best Local Similarity   50.0%;  Pred. No. 6e-10;
  Matches   22;  Conservative    6;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              | :|:  |: ||| | ||   |||  |  |||  :||| ||:|:
Db         49 AFNFMEELKPGSLERECIEEKCDFEEAFEIFETKEDTLNFWAKY 92


RESULT 8
Q8N2N6
ID   Q8N2N6      PRELIMINARY;      PRT;   231 AA.
AC   Q8N2N6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein FLJ90093.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA   Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA   Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA   Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK074574; BAC11069.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   231 AA;  25844 MW;  8A373B0D5C1D0D81 CRC64;

  Query Match             54.3%;  Score 107;  DB 4;  Length 231;
  Best Local Similarity   43.9%;  Pred. No. 6.3e-10;
  Matches   18;  Conservative    8;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  ||||:: | |:  || :   | :||: : |: ||
Db         20 ANEFLEELRQGTIERECMEEICSYEEVKEVFENKEKTMEFW 60


RESULT 9
Q804X5
ID   Q804X5      PRELIMINARY;      PRT;   433 AA.
AC   Q804X5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Anticoagulant protein C precursor (EC 3.4.21.69).
GN   PROC.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
```

```
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RT   "Comparative sequence analysis and molecular evolution of blood
RT   coagulation genes from Gallus gallus and Fugu rubripes.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF465270; AAO33365.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003808; F:protein C (activated) activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase.
SQ   SEQUENCE   433 AA;  48689 MW;  E09DDEE56D7DA2A3 CRC64;

  Query Match             54.3%;  Score 107;  DB 13;  Length 433;
  Best Local Similarity   50.0%;  Pred. No. 1.2e-09;
  Matches   22;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |: ||: | |    |:|  |  |||  : || ||||:
Db         40 ANSFLEELKPGSVERECNEERCNFEEASEIFETKEATLEFWSKY 83


RESULT 10
Q7SXH8
ID   Q7SXH8      PRELIMINARY;      PRT;   524 AA.
AC   Q7SXH8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC055596; AAH55596.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   524 AA;  59560 MW;  1B4AE08119080325 CRC64;

  Query Match             53.3%;  Score 105;  DB 13;  Length 524;
  Best Local Similarity   43.2%;  Pred. No. 3.4e-09;
  Matches   19;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||:    |: |:| | |:  |||   |: :|| || |  ||:|:
Db         42 ANTVFEELKPGNLERECVEEICDHEEAREVFERVDKTEIFWAKY 85


RESULT 11
Q8JHC9
ID   Q8JHC9      PRELIMINARY;      PRT;   443 AA.
AC   Q8JHC9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor VIIi.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hanumanthaiah R., Day K., Jagadeeswaran P.;
RT   "Comprehensive analysis of blood coagulation pathways in Teleostei:
RT   Evolution of coagulation factor genes and identification of zebrafish
RT   factor VIIi.";
RL   Blood Cells Mol. Dis. 0:0-0(2002).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF519546; AAM88342.1; -.
DR   EMBL; AF515269; AAN71000.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
```

```
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   443 AA;  48823 MW;  2D2504718AE94FF4 CRC64;

  Query Match             52.5%;  Score 103.5;  DB 13;  Length 443;
  Best Local Similarity   47.6%;  Pred. No. 5.1e-09;
  Matches   20;  Conservative    7;  Mismatches   14;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||| ||  :: |:| | |:  |||:  |: :||| | |  ||
Db         38 ANSGFLEEMKAGNLERECVEEICDYEEAREVFEDDDRTKQFW 79


RESULT 12
Q80Y26
ID   Q80Y26      PRELIMINARY;      PRT;   340 AA.
AC   Q80Y26;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   F10 protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22386257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RA   Strausberg R.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC050219; AAH50219.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
SQ   SEQUENCE   340 AA;  38359 MW;  EE252D6157720811 CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 340;
  Best Local Similarity   38.6%;  Pred. No. 1e-08;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||    ::|:| | |:  || :   : |||| : |  :|:|:
Db         53 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 96


RESULT 13
O54740
ID   O54740      PRELIMINARY;      PRT;   481 AA.
AC   O54740;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6).
GN   F10 OR FA10.
OS   Mus musculus (Mouse).
OG   Plasmid pBluescript.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=98454993; PubMed=9783672;
RA   Heidtmann H.H., Kontermann R.E.;
RT   "Cloning and recombinant expression of mouse coagulation factor X.";
RL   Thromb. Res. 92:33-41(1998).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AJ222677; CAA10933.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0003804; F:blood coagulation factor X activity; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
```

```
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease; Signal;
KW   Plasmid.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41    481       COAGULATION FACTOR X.
SQ   SEQUENCE   481 AA;  53986 MW;  CF702DE5EF9D97AE CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 1.5e-08;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||    ::|:| | |:  || :   : |||| : |  :|:|:
Db         41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 14
Q99L32
ID   Q99L32      PRELIMINARY;      PRT;   481 AA.
AC   Q99L32;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor X.
GN   F10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; BC003877; AAH03877.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   481 AA;  54004 MW;  BD88E96C8A0B7E7F CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 1.5e-08;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||    ::|:| | |:  || :   : |||| : |  :|:|:
Db         41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 15
O88947
ID   O88947      PRELIMINARY;      PRT;   481 AA.
AC   O88947;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor X precursor.
GN   F10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL6 X CBA; TISSUE=Liver;
RX   MEDLINE=98347933; PubMed=9684791;
RA   Liang Z., Cooper A., DeFord M.E., Carmeliet P., Collen D.,
RA   Castellino F.J., Rosen E.D.;
RT   "Cloning and characterization of a cDNA encoding murine coagulation
RT   factor X.";
RL   Thromb. Haemost. 80:87-91(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SJ;
RA   Cooper A., Liang Z., Castellino F.J., Rosen E.D.;
RT   "Cloning and Characterization of the Murine Factor X Gene.";
RL   Thromb. Haemost. 0:0-0(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF087644; AAC36345.1; -.
DR   EMBL; AF211347; AAF22980.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
```

```
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease; Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41    481       COAGULATION FACTOR X.
SQ   SEQUENCE   481 AA;  54018 MW;  8AC09DE5EF9D271E CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 1.5e-08;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||     ::|:| | |:  || :    : |||| : |  :|:|:
Db         41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


Search completed: March  1, 2004, 10:02:52
Job time : 38.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 10:01:28 ; Search time 28 Seconds
(without alignments)
331.812 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	419	14	US-10-182-263-6	Sequence 6, Appli
2	179	90.9	419	15	US-10-168-407-5	Sequence 5, Appli
3	179	90.9	419	15	US-10-168-407-6	Sequence 6, Appli
4	174	88.3	419	14	US-10-182-263-3	Sequence 3, Appli
5	174	88.3	419	14	US-10-182-263-4	Sequence 4, Appli
6	174	88.3	419	14	US-10-182-263-5	Sequence 5, Appli
7	174	88.3	419	15	US-10-168-407-3	Sequence 3, Appli
8	174	88.3	419	15	US-10-168-407-4	Sequence 4, Appli
9	160	81.2	44	14	US-10-298-330-1	Sequence 1, Appli
10	160	81.2	419	10	US-09-978-917A-4	Sequence 4, Appli
11	160	81.2	419	14	US-10-182-263-1	Sequence 1, Appli
12	160	81.2	419	15	US-10-168-407-1	Sequence 1, Appli
13	160	81.2	461	10	US-09-978-917A-2	Sequence 2, Appli
14	160	81.2	461	14	US-10-182-263-2	Sequence 2, Appli
15	160	81.2	461	15	US-10-168-407-2	Sequence 2, Appli
16	116	58.9	44	14	US-10-298-330-2	Sequence 2, Appli
17	110	55.8	139	15	US-10-360-101-232	Sequence 232, App
18	110	55.8	488	14	US-10-348-504-44	Sequence 44, Appl
19	110	55.8	488	14	US-10-407-123-27	Sequence 27, Appl
20	106	53.8	44	14	US-10-298-330-18	Sequence 18, Appl
21	99	50.3	42	16	US-10-038-854-97	Sequence 97, Appl
22	99	50.3	406	10	US-09-782-587B-3	Sequence 3, Appli
23	99	50.3	406	15	US-10-383-898-1	Sequence 1, Appli
24	99	50.3	466	14	US-10-017-122-2	Sequence 2, Appli
25	99	50.3	466	15	US-10-375-741-14	Sequence 14, Appl
26	96	48.7	44	14	US-10-298-330-3	Sequence 3, Appli
27	96	48.7	406	10	US-09-782-587B-1	Sequence 1, Appli
28	96	48.7	406	14	US-10-109-498-1	Sequence 1, Appli
29	96	48.7	406	14	US-10-255-032-1	Sequence 1, Appli
30	96	48.7	406	14	US-10-281-727-1	Sequence 1, Appli
31	96	48.7	406	15	US-10-386-898-7	Sequence 7, Appli
32	95	48.2	405	15	US-10-360-101-225	Sequence 225, App
33	91	46.2	40	14	US-10-298-330-23	Sequence 23, Appl
34	90	45.7	44	14	US-10-298-330-4	Sequence 4, Appli
35	89	45.2	40	14	US-10-298-330-25	Sequence 25, Appl
36	85	43.1	40	14	US-10-298-330-22	Sequence 22, Appl
37	84.5	42.9	96	10	US-09-759-130B-313	Sequence 313, App
38	84.5	42.9	96	14	US-10-189-123-43	Sequence 43, Appl
39	84.5	42.9	96	14	US-10-188-495-43	Sequence 43, Appl
40	84.5	42.9	209	10	US-09-759-130B-312	Sequence 312, App
41	84.5	42.9	209	14	US-10-189-123-42	Sequence 42, Appl
42	84.5	42.9	209	14	US-10-188-495-42	Sequence 42, Appl
43	84.5	42.9	226	10	US-09-759-130B-310	Sequence 310, App
44	84.5	42.9	226	14	US-10-189-123-40	Sequence 40, Appl
45	84.5	42.9	226	14	US-10-188-495-40	Sequence 40, Appl

ALIGNMENTS

PATENT 6630138 SILVER, DAN SWOPE PRIORITY 2-11-00

RESULT 1
US-10-182-263-6
; Sequence 6, Application US/10182263
; Publication No. US20030023554A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-6

Query Match 90.9%; Score 179; DB 14; Length 419;
Best Local Similarity 79.5%; Pred. No. 3.9e-22;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |||||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44
```

RESULT 2
US-10-168-407-5
; Sequence 5, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:

```
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-5

  Query Match             90.9%;  Score 179;  DB 15;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 3.9e-22;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  || ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 3
US-10-168-407-6
; Sequence 6, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-6

  Query Match             90.9%;  Score 179;  DB 15;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 3.9e-22;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  || ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 4
US-10-182-263-3
; Sequence 3, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-3

  Query Match             88.3%;  Score 174;  DB 14;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 2.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  |||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 5
US-10-182-263-4
; Sequence 4, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-4

  Query Match             88.3%;  Score 174;  DB 14;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 2.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  |||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 6
US-10-182-263-5
; Sequence 5, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-5

  Query Match             88.3%;  Score 174;  DB 14;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 2.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
                  ||||| || ||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 7
US-10-168-407-3
; Sequence 3, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-3

  Query Match             88.3%;  Score 174;  DB 15;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 2.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||| || ||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 8
US-10-168-407-4
; Sequence 4, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-4

  Query Match             88.3%;  Score 174;  DB 15;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 2.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||| || ||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 9
US-10-298-330-1
; Sequence 1, Application US/10298330
; Publication No. US20030100506A1
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary L.
;  TITLE OF INVENTION: Modified Vitamin K-Dependent
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 09531-127001
;  CURRENT APPLICATION NUMBER: US/10/298,330
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: 09/497,591
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: 09/302,239
;  PRIOR FILING DATE: 1999-04-29
;  PRIOR APPLICATION NUMBER: 08/955,636
;  PRIOR FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 6, 7, 14, 16, 19, 20, 25, 26, 29
;   OTHER INFORMATION: Xaa = gamma carboxyglutamic or glutamic acid
US-10-298-330-1

  Query Match             81.2%;  Score 160;  DB 14;  Length 44;
  Best Local Similarity   90.9%;  Pred. No. 5.7e-20;
  Matches   40;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||||::||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 10
US-09-978-917A-4
; Sequence 4, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us310 - protein C
;  CURRENT APPLICATION NUMBER: US/09/978,917A
;  CURRENT FILING DATE:  2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-978-917A-4

  Query Match             81.2%;  Score 160;  DB 10;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 7e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||| ||  || | ||  |||| || ||::||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 11
US-10-182-263-1
; Sequence 1, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 419
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-10-182-263-1

  Query Match             81.2%;  Score 160;  DB 14;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 7e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||::||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 12
US-10-168-407-1
; Sequence 1, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-1

  Query Match             81.2%;  Score 160;  DB 15;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 7e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||::||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 13
US-09-978-917A-2
; Sequence 2, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us310 - protein C
;  CURRENT APPLICATION NUMBER: US/09/978,917A
;  CURRENT FILING DATE:  2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: (1)...(42)
;   FEATURE:
;   NAME/KEY: CHAIN
;   LOCATION: (43)...(461)
US-09-978-917A-2

  Query Match             81.2%;  Score 160;  DB 10;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 7.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||::||||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 14
US-10-182-263-2
; Sequence 2, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-2

  Query Match             81.2%;  Score 160;  DB 14;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 7.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||::||||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 15
US-10-168-407-2
; Sequence 2, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-2

  Query Match             81.2%;  Score 160;  DB 15;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 7.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||::||||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


Search completed: March  1, 2004, 10:11:10
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:54:37 ; Search time 50 Seconds
(without alignments)
248.642 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	180	90.9	44	2	AAY18297	Aay18297 Modified
2	176	88.9	419	4	AAB82676	Aab82676 Human pro
3	176	88.9	419	4	AAB82675	Aab82675 Human pro
4	176	88.9	419	4	AAE08628	Aae08628 Human pro
5	176	88.9	419	4	AAE08627	Aae08627 Human pro
6	176	88.9	419	4	AAE08629	Aae08629 Human pro
7	174	87.9	45	5	ABB79950	Abb79950 Human pro
8	172	86.9	44	2	AAY18299	Aay18299 Modified
9	172	86.9	44	2	AAY18309	Aay18309 Modified
10	172	86.9	44	2	AAY18298	Aay18298 Modified
11	172	86.9	44	2	AAY18303	Aay18303 Human pro
12	172	86.9	44	4	AAB36402	Aab36402 Human pro
13	172	86.9	44	7	ADD50094	Add50094 Human vit
14	172	86.9	45	2	AAW75710	Aaw75710 Partial h
15	172	86.9	45	5	ABB79947	Abb79947 Human pro
16	172	86.9	45	7	ADB71155	Adb71155 Human pro
17	172	86.9	415	3	AAY56803	Aay56803 Truncated
18	172	86.9	419	2	AAR35760	Aar35760 Protein C
19	172	86.9	419	2	AAW72753	Aaw72753 Primary s
20	172	86.9	419	4	AAB82673	Aab82673 Wild-type
21	172	86.9	419	4	AAB36896	Aab36896 Human pro
22	172	86.9	419	4	AAB36897	Aab36897 Human pro
23	172	86.9	419	4	AAB36898	Aab36898 Human pro
24	172	86.9	419	4	AAB36894	Aab36894 Human pro
25	172	86.9	419	4	AAE08625	Aae08625 Human mat
26	172	86.9	419	5	AAU99005	Aau99005 Human Pro
27	172	86.9	419	5	AAU99006	Aau99006 Human Pro
28	172	86.9	419	5	AAU99008	Aau99008 Human Pro
29	172	86.9	419	5	AAU99018	Aau99018 Human Pro
30	172	86.9	419	5	AAU99026	Aau99026 Human Pro
31	172	86.9	419	5	AAU99037	Aau99037 Human Pro
32	172	86.9	419	5	AAU99049	Aau99049 Human Pro
33	172	86.9	419	5	AAU99063	Aau99063 Human Pro
34	172	86.9	419	5	AAU99072	Aau99072 Human Pro
35	172	86.9	419	5	AAU99083	Aau99083 Human Pro
36	172	86.9	419	5	AAU99084	Aau99084 Human Pro
37	172	86.9	419	5	AAU99012	Aau99012 Human Pro
38	172	86.9	419	5	AAU99021	Aau99021 Human Pro
39	172	86.9	419	5	AAU99033	Aau99033 Human Pro
40	172	86.9	419	5	AAU99039	Aau99039 Human Pro
41	172	86.9	419	5	AAU99047	Aau99047 Human Pro
42	172	86.9	419	5	AAU99050	Aau99050 Human Pro
43	172	86.9	419	5	AAU99076	Aau99076 Human Pro
44	172	86.9	419	5	AAU99088	Aau99088 Human Pro
45	172	86.9	419	5	AAU99097	Aau99097 Human Pro

ALIGNMENTS

RESULT 1
AAY18297
ID AAY18297 standard; peptide; 44 AA.
XX
AC AAY18297;
XX
DT 17-AUG-1999 (first entry)
XX
DE Modified GLA domain of vitamin K-dependent protein.
XX
KW GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .44
FT /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT acid"
XX
PN WO9920767-A1.
XX
PD 29-APR-1999.
XX
PF 20-OCT-1998; 98WO-US022152.
XX
PR 23-OCT-1997; 97US-00955636.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Nelsestuen GL;
XX
DR WPI; 1999-288309/24.
XX
PT Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT domain, useful for treating clotting disorders.
XX
PS Claim 6; Page 78; 86pp; English.
XX
CC This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC domain. The invention relates to a vitamin K-dependent polypeptide
CC comprising a modified GLA domain containing an amino acid substitution
CC which enhances membrane binding of the modified polypeptide as compared
CC to the native polypeptide. The polypeptide is used to treat a clotting
CC disorder by decreasing or increasing clot formation. Modification of the
CC GLA domain results in a protein which has enhanced membrane binding
CC affinity as compared to the native protein

```
XX
SQ   Sequence 44 AA;

  Query Match             90.9%;  Score 180;  DB 2;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-22;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 2
AAB82676
ID   AAB82676 standard; protein; 419 AA.
XX
AC   AAB82676;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S/T254S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Misc-difference 254
FT                   /note= "Thr in wild-type protein"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26364.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 4; Page 53-54; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which Ser at position 11 of the mature wild-type protein C sequence (see
CC   AAB82673) is substituted with Gly, Gln at position 32 with Glu, Asn at
CC   position 33 with Asp, Leu at position 194 with Ser, and Thr at position
CC   254 with Ser. It is an example of protein C derivatives of the invention
CC   that have at least 2 amino acid substitutions, but which have increased
CC   anticoagulant activity and resistance to inactivation by serpins compared
CC   with the wild-type protein, while retaining the biological activity of
CC   the wild-type protein. A method of producing the derivatives using
CC   recombinant DNA methods is claimed. The protein C derivatives are useful
CC   for treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina), vascular
CC   occlusive disorders and hypercoagulable states, sepsis (in combination
CC   with bactericidal permeability increasing protein or with tissue factor
CC   pathway inhibitor), thrombotic disorders (in combination with an anti-
CC   platelet agent or by local delivery through an intracoronary catheter),
CC   protein C deficiency, acute arterial thrombotic occlusion,
CC   thromboembolism, or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts. Human patients with genetically predisposed
CC   prothrombotic disorders may be treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             88.9%;  Score 176;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.9e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||| || | ||  ||||  || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 3
AAB82675
ID   AAB82675 standard; protein; 419 AA.
XX
AC   AAB82675;
```

```
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26363.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 3; Page 52-53; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which Ser at amino acid position 11 of the mature wild-type protein C
CC   sequence (see AAB82673) is substituted with Gly, Gln at position 32 with
CC   Glu, Asn at position 33 with Asp, and Leu at position 194 with Ser. The
CC   protein is an example of protein C derivatives of the invention that have
CC   at least 2 amino acid substitutions, but which have increased
CC   anticoagulant activity and resistance to inactivation by serpins compared
CC   with the wild-type protein, while retaining the biological activity of
CC   the wild-type protein. A method of producing the derivatives using
CC   recombinant DNA methods is claimed. The protein C derivatives are useful
CC   for treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina), vascular
CC   occlusive disorders and hypercoagulable states, sepsis (in combination
CC   with bactericidal permeability increasing protein or with tissue factor
CC   pathway inhibitor), thrombotic disorders (in combination with an anti-
CC   platelet agent or by local delivery through an intracoronary catheter),
CC   protein C deficiency, acute arterial thrombotic occlusion,
CC   thromboembolism, or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts. Human patients with genetically predisposed
CC   prothrombotic disorders may be treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             88.9%;  Score 176;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.9e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||| || | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 4
AAE08628
ID   AAE08628 standard; protein; 419 AA.
XX
AC   AAE08628;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #2.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
```

PF 02-FEB-2001; 2001WO-US001221.
XX
PR 11-FEB-2000; 2000US-0181948P.
PR 14-MAR-2000; 2000US-0189199P.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Gerlitz BE, Grinnell BW, Jones BE;
XX
DR WPI; 2001-514662/56.
DR N-PSDB; AAD15226.
XX
PT Protein C derivative for treating acute coronary syndromes, vascular
PT occlusive disorders, thrombotic disorders and sepsis, comprises
PT substitutions at specified amino acid positions.
XX
PS Claim 4; Page 47-48; 59pp; English.
XX
CC The invention relates to human protein C derivatives and nucleic acid
CC molecules encoding such derivatives. These derivatives have increased
CC anticoagulation activity, resistance to serpin inactivation and increased
CC sensitivity to thrombin activation compared to wild type protein C, and
CC retains the biological activity of the wild type human protein C. Protein
CC C derivatives are useful in the manufacture of a medicament for the
CC treatment of acute coronary syndromes e.g. myocardial infarction and
CC unstable angina; and disease states predisposing to thrombosis; vascular
CC occlusive disorders and hypercoagulable states e.g. disseminated
CC intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC syndrome; sepsis in combination with bacterial permeability increasing
CC protein; thrombotic disorders in combination with an anti-platelet agent;
CC protein C deficiency; acute arterial thrombotic occlusion,
CC thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC acid molecules of the invention are useful for treating humans with
CC genetically predisposed prothrombotic disorders by gene therapy. The
CC present sequence is human protein C derivative
XX
SQ Sequence 419 AA;

Query Match 88.9%; Score 176; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 1.9e-20;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 5
AAE08627
ID AAE08627 standard; protein; 419 AA.
XX
AC AAE08627;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human protein C derivative #1.
XX
KW Human; protein C derivative; anticoagulation activity; thrombosis;
KW serpin inactivation; acute coronary syndrome; myocardial infarction;
KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW disseminated intravascular coagulation; DIC; burn; transplantation;
KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS Homo sapiens.
XX
PN WO200159084-A1.
XX
PD 16-AUG-2001.
XX

PF 02-FEB-2001; 2001WO-US001221.
XX
PR 11-FEB-2000; 2000US-0181948P.
PR 14-MAR-2000; 2000US-0189199P.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Gerlitz BE, Grinnell BW, Jones BE;
XX
DR WPI; 2001-514662/56.
DR N-PSDB; AAD15225.
XX
PT Protein C derivative for treating acute coronary syndromes, vascular
PT occlusive disorders, thrombotic disorders and sepsis, comprises
PT substitutions at specified amino acid positions.
XX
PS Claim 3; Page 46-47; 59pp; English.
XX
CC The invention relates to human protein C derivatives and nucleic acid
CC molecules encoding such derivatives. These derivatives have increased
CC anticoagulation activity, resistance to serpin inactivation and increased
CC sensitivity to thrombin activation compared to wild type protein C, and
CC retains the biological activity of the wild type human protein C. Protein
CC C derivatives are useful in the manufacture of a medicament for the
CC treatment of acute coronary syndromes e.g. myocardial infarction and
CC unstable angina; and disease states predisposing to thrombosis; vascular
CC occlusive disorders and hypercoagulable states e.g. disseminated
CC intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC syndrome; sepsis in combination with bacterial permeability increasing
CC protein; thrombotic disorders in combination with an anti-platelet agent;
CC protein C deficiency; acute arterial thrombotic occlusion,
CC thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC acid molecules of the invention are useful for treating humans with
CC genetically predisposed prothrombotic disorders by gene therapy. The
CC present sequence is human protein C derivative
XX
SQ Sequence 419 AA;

Query Match 88.9%; Score 176; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 1.9e-20;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 6
AAE08629
ID AAE08629 standard; protein; 419 AA.
XX
AC AAE08629;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human protein C derivative #3.
XX
KW Human; protein C derivative; anticoagulation activity; thrombosis;
KW serpin inactivation; acute coronary syndrome; myocardial infarction;
KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW disseminated intravascular coagulation; DIC; burn; transplantation;
KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "Encoded by CAA"
XX

```
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US001221.
XX
PR   11-FEB-2000; 2000US-0181948P.
PR   14-MAR-2000; 2000US-0189199P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15227.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions.
XX
PS   Claim 5; Page 48-49; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and increased
CC   sensitivity to thrombin activation compared to wild type protein C, and
CC   retains the biological activity of the wild type human protein C. Protein
CC   C derivatives are useful in the manufacture of a medicament for the
CC   treatment of acute coronary syndromes e.g. myocardial infarction and
CC   unstable angina; and disease states predisposing to thrombosis; vascular
CC   occlusive disorders and hypercoagulable states e.g. disseminated
CC   intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC   sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC   syndrome; sepsis in combination with bacterial permeability increasing
CC   protein; thrombotic disorders in combination with an anti-platelet agent;
CC   protein C deficiency; acute arterial thrombotic occlusion,
CC   thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC   acid molecules of the invention are useful for treating humans with
CC   genetically predisposed prothrombotic disorders by gene therapy. The
CC   present sequence is human protein C derivative
XX
SQ   Sequence 419 AA;

  Query Match             88.9%;  Score 176;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.9e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||| || | ||  ||||  || ||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 7
ABB79950
ID   ABB79950 standard; protein; 45 AA.
XX
AC   ABB79950;
XX
DT   12-DEC-2002  (first entry)
XX
DE   Human protein C mutated Gla domain SED.
XX
KW   Protein C; Gla domain; human; blood clotting; anticoagulant;
KW   thrombolytic; antiarteriosclerotic; cardiant; antiaggregant; mutant;
KW   mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 23
FT                   /note= "wild-type Asp substituted by Ser"
FT   Misc-difference 32
FT                   /note= "wild-type Gln substituted by Glu"
FT   Misc-difference 33
FT                   /note= "wild-type Asn substituted by Asp"
XX
PN   WO200270681-A1.
XX
PD   12-SEP-2002.
XX
PF   01-MAR-2002; 2002WO-SE000363.
XX
PR   02-MAR-2001; 2001US-0272466P.
XX
PA   (TACT-) TAC THROMBOSIS & COAGULATION AB.
XX
PI   Dahlbaeck B;
XX
DR   WPI; 2002-713449/77.
XX
PT   New variant blood coagulation component, useful for manufacturing a
PT   medicament for treating or preventing coagulation disorders, e.g.
PT   thrombosis, comprises an anticoagulant activity in the protein C-
PT   anticoagulant system of blood.
XX
PS   Example 1; Page; 58pp; English.
XX
CC   The present sequence is the protein sequence of a mutated Gla domain (N-
CC   terminal amino acids 1-45) of human protein C. The mutated Gla domain
CC   contains the substitution mutations D23S, Q32E and N33D. Protein C and
CC   activated protein C variants comprising a mutated Gla domain are provided
CC   by the invention. The variants contain at least 6, and optionally 7-10,
CC   amino acid substitutions. A preferred mutant (designated QGNSEDY, see
CC   ABB79946) has the mutations H10Q, S11G, S12N, D23S, Q32E, N33D and H44Y,
CC   and shows greatly enhanced anticoagulant activity in standard in vitro
CC   coagulation assays. The present mutant (designated SED) was produced in
CC   an example from the invention as a step toward the production of the
CC   QGNSEDY mutant Gla domain, and shows little, if any, improvement in
CC   anticoagulant activity over wild-type activated protein C. The invention
CC   provides methods for producing the variants based on DNA technology, and
CC   with the use of the variants for the treatment of coagulation disorders
CC   such as thrombosis or APC resistance, or in diagnostic test systems for
CC   assaying components of the protein C-anticoagulant system (all claimed).
CC   The variants may also be used in treating arteriosclerosis, myocardial
CC   infarction, and disseminated intravascular coagulation. Note: The present
CC   sequence is not shown in the specification but is derived from the human
CC   wild-type Gla domain sequence given on page 7 of the specification (see
CC   ABB79947)
XX
SQ   Sequence 45 AA;

  Query Match             87.9%;  Score 174;  DB 5;  Length 45;
  Best Local Similarity   77.3%;  Pred. No. 3.6e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |||||| | ||  || |  || ||||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICSFEEAKEIFEDVDDTLAFWSKH 44


RESULT 8
AAY18299
ID   AAY18299 standard; peptide; 44 AA.
XX
AC   AAY18299;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 8; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 2;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 7.6e-21;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||||||| |||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 9
AAY18309
ID   AAY18309 standard; peptide; 44 AA.
XX
AC   AAY18309;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Disclosure; Page 79-80; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 2;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 7.6e-21;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||||||||||||||||||||||||| |||||||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAFXIFEDVDDTLAFWSKH 44


RESULT 10
AAY18298
ID   AAY18298 standard; peptide; 44 AA.
XX
AC   AAY18298;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 7; Page 78; 86pp; English.
XX
```

```
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 2;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 7.6e-21;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||||||| ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 11
AAY18303
ID   AAY18303 standard; peptide; 44 AA.
XX
AC   AAY18303;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Human protein C GLA domain.
XX
KW   GLA domain; vitamin K-dependent protein; clotting disorder; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Disclosure; Page 14; 86pp; English.
XX
CC   This sequence is the protein C GLA (gamma-carboxyglutamic acid) domain.
CC   The invention relates to a vitamin K-dependent polypeptide comprising a
CC   modified GLA domain containing an amino acid substitution which enhances
CC   membrane binding of the modified polypeptide as compared to the native
CC   polypeptide. The polypeptide is used to treat a clotting disorder by
CC   decreasing or increasing clot formation. Modification of the GLA domain
CC   results in a protein which has enhanced membrane binding affinity as
CC   compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 2;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-21;
  Matches   42;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||||||||||||||||||||||::|||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 12
AAB36402
ID   AAB36402 standard; peptide; 44 AA.
XX
AC   AAB36402;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human protein C gamma-carboxyglutamic acid domain SEQ ID NO:1.
XX
KW   Vitamin K-dependent protein; factor VII; protein C; GLA domain;
KW   gamma-carboxyglutamic acid domain; factor IX; protein S; protein Z;
KW   factor X; prothrombin; enhanced membrane binding affinity;
KW   clot formation; thrombolytic; haemostatic; bleeding disorder; thrombosis;
KW   clotting disorder; haemophilia A; haemophilia B; liver disease.
XX
OS   Homo sapiens.
XX
PN   WO200066753-A2.
XX
PD   09-NOV-2000.
XX
PF   28-APR-2000; 2000WO-US011416.
XX
PR   29-APR-1999;   99US-00302239.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 2001-007226/01.
XX
PT   Novel vitamin K-dependent polypeptide useful for treating clotting
PT   disorders such as thrombosis and hemophilia, comprises modified gamma-
PT   carboxy glutamic acid domain that enhances membrane binding affinity.
XX
PS   Example 5; Page 42; 81pp; English.
XX
CC   The present invention describes a vitamin K-dependent polypeptide (I)
CC   comprising a modified gamma-carboxy glutamic acid (GLA) domain having at
CC   least one amino acid substitution, that enhances membrane binding
CC   affinity and the activity of the polypeptide relative to a corresponding
CC   native vitamin K-dependent polypeptide and inhibits clot formation. (I)
CC   can have thrombolytic and haemostatic activities, and can be used as an
CC   inhibitor of clot formation. (I) is useful for decreasing clot formation
CC   in a mammal, a factor VII or factor IX containing a modified GLA domain
CC   is useful for increasing clot formation and for treating a bleeding
CC   disorder, including thrombosis and clotting disorders such as haemophilia
CC   A, haemophilia B and liver disease. The present sequence represents a
CC   human protein C GLA domain sequence, given in the exemplification of the
CC   present invention
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 4;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-21;
  Matches   42;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||||||||||||||||||||||::|||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 13
ADD50094
ID   ADD50094 standard; protein; 44 AA.
XX
AC   ADD50094;
```

```
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human vitamin K-dependent protein #1.
XX
KW   Human; vitamin K-dependent protein; gamma-carboxyglutamic acid domain;
KW   GLA domain; membrane binding affinity; clot formation; haemophilia;
KW   clotting disorder; site directed mutagenesis; haemostatic; anticoagulant;
KW   thrombolytic.
XX
OS   Homo sapiens.
XX
PN   US2003100506-A1.
XX
PD   29-MAY-2003.
XX
PF   18-NOV-2002; 2002US-00298330.
XX
PR   23-OCT-1997;   97US-00955636.
PR   29-APR-1999;   99US-00302239.
PR   03-FEB-2000; 2000US-00497591.
XX
PA   (NELS/) NELSESTUEN G L.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 2003-606646/57.
XX
PT   New vitamin K-dependent polypeptide for modulating clot formation in
PT   mammals comprises a modified gamma-carboxyglutamic acid domain that
PT   enhances membrane binding affinity and activity of the polypeptide.
XX
PS   Example 5; SEQ ID NO 1; 51pp; English.
XX
CC   The invention relates to a vitamin K-dependent polypeptide comprising a
CC   modified gamma-carboxyglutamic acid (GLA) domain that enhances membrane
CC   binding affinity and activity of the polypeptide relative to a
CC   corresponding native vitamin K-dependent polypeptide, where the modified
CC   GLA domain comprises a glutamic acid residue at position 34. The
CC   polypeptide is useful in modulating clot formation in mammals or in
CC   treating certain types of haemophilia or clotting disorders. The membrane
CC   binding affinity of polypeptides is increased by site directed
CC   mutagenesis in the GLA domain. This sequence represents a vitamin K-
CC   dependent protein of the invention.
XX
SQ   Sequence 44 AA;

  Query Match             86.9%;  Score 172;  DB 7;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-21;
  Matches   42;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||||||||||||||||||||||::|||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 14
AAW75710
ID   AAW75710 standard; protein; 45 AA.
XX
AC   AAW75710;
XX
DT   08-DEC-1998  (first entry)
XX
DE   Partial human protein C amino acid sequence.
XX
KW   Gamma carboxylglutamic acid; human protein C; Gla domain; chimera;
KW   pRc/RSV; RSV-PC; amplification, PCR, primer; transfection; anticoagulant;
KW   human 293 cell; Protein S; myocardial infarction; venous thrombosis;
KW   disseminated intravascular coagulation; thromboembolic disease; lupus;
KW   adult respiratory distress syndrome; factor V Leiden; stroke.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 6
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 7
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 14
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 16
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 19
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 20
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 25
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 26
FT                   /note= "Gamma carboxylglutamic acid"
FT   Misc-difference 29
FT                   /note= "Gamma carboxylglutamic acid"
XX
PN   WO9820118-A1.
XX
PD   14-MAY-1998.
XX
PF   07-NOV-1997;   97WO-US020376.
XX
PR   08-NOV-1996;   96US-00745254.
PR   25-JUL-1997;   97US-0053768P.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   Esmon CT,  Smirnov M;
XX
DR   WPI; 1998-286934/25.
XX
PT   Protein C chimeric proteins for use as anticoagulants - having gamma
PT   carboxyglutamic acid region replaced with Vitamin K dependent clotting
PT   factor e.g. prothrombin.
XX
PS   Example 1; Page 15; 42pp; English.
XX
CC   The present sequence represents the first three exons of the human
CC   protein C protein, which contains gamma carboxylglutamic acid modified
CC   residues. This sequence was replaced with the corresponding regions of
CC   modified human prothrombin (AAW75709), to create a protein C prothrombin
CC   Gla domain chimera. To produce this chimera, the wild-type protein C cDNA
CC   was ligated into pRc/RSV to form RSV-PC, and then was digested with
CC   restriction enzymes to remove the first three exons and the first codon
CC   of exon four. The prothrombin cDNA was amplified, digested, and then
CC   ligated into RSV-PC at the identical site where the protein C cDNA exons
CC   1-3 had been removed. This construct was then transfected into human 293
CC   cells, from which the chimeric protein can be purified. This chimeric
CC   protein can be used as a anticoagulant, to treat disorders where protein
CC   S is low, some forms of lupus, following stroke or myocardial infarction,
CC   after venous thrombosis and in disseminated intravascular coagulation,
CC   adult respiratory distress syndrome, in thromboembolic disease or factor
CC   V Leiden
XX
SQ   Sequence 45 AA;

  Query Match             86.9%;  Score 172;  DB 2;  Length 45;
  Best Local Similarity   95.5%;  Pred. No. 7.8e-21;
  Matches   42;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||||||||||||||||||||||::|||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 15
ABB79947
ID   ABB79947 standard; protein; 45 AA.
```

XX
AC ABB79947;
XX
DT 12-DEC-2002 (first entry)
XX
DE Human protein C Gla domain.
XX
KW Protein C; Gla domain; human; blood clotting; anticoagulant;
KW thrombolytic; antiarteriosclerotic; cardiant; antiaggregant.
XX
OS Homo sapiens.
XX
PN WO200270681-A1.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-SE000363.
XX
PR 02-MAR-2001; 2001US-0272466P.
XX
PA (TACT-) TAC THROMBOSIS & COAGULATION AB.
XX
PI Dahlbaeck B;
XX
DR WPI; 2002-713449/77.
XX
PT New variant blood coagulation component, useful for manufacturing a
PT medicament for treating or preventing coagulation disorders, e.g.
PT thrombosis, comprises an anticoagulant activity in the protein C-
PT anticoagulant system of blood.
XX
PS Disclosure; Page 7; 58pp; English.
XX
CC The present sequence is the protein sequence of the Gla domain (N-
CC terminal amino acids 1-45) of human protein C. The invention provides
CC human protein C and activated protein C variants in which the Gla domain
CC contains at least 6, and optionally 7-10, amino acid substitutions,
CC preferably the substitution mutations H10Q, S11G, S12N, D23S, Q32E, N33D
CC and H44Y (see ABB79946). Protein C variants comprising the mutated Gla
CC domain show much enhanced anticoagulant activity, as shown in increased
CC clotting time in standard in vitro coagulation assays, as a result of
CC enhanced calcium and/or membrane binding properties. The invention
CC provides methods for producing the variants based on DNA technology, and
CC for using the variants in the treatment of coagulation disorders such as
CC thrombosis or APC resistance, or in diagnostic test systems for assaying
CC components of the protein C-anticoagulant system (all claimed). The
CC variants may also be used in treating arteriosclerosis, myocardial
CC infarction, and disseminated intravascular coagulation
XX
SQ Sequence 45 AA;

Query Match 86.9%; Score 172; DB 5; Length 45;
Best Local Similarity 75.0%; Pred. No. 7.8e-21;
Matches 33; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

Search completed: March 1, 2004, 10:01:23
Job time : 52 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:59:33 ; Search time 13.5 Seconds
(without alignments)
313.513 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	172	86.9	461	1	KXHU	protein C (activat
2	133	67.2	461	1	JX0210	protein C (activat
3	132	66.7	461	1	S18994	protein C (activat
4	115	58.1	456	1	KXBO	protein C (activat
5	107	54.0	482	1	EXRT	coagulation factor
6	103	52.0	492	1	EXBO	coagulation factor
7	102	51.5	488	1	EXHU	coagulation factor
8	94	47.5	443	2	I46932	coagulation factor
9	92	46.5	466	1	KFHU7	coagulation factor
10	78.5	39.6	617	2	S10511	thrombin (EC 3.4.2
11	78.5	39.6	618	2	A35827	thrombin (EC 3.4.2
12	78	39.4	407	1	KFBO7	coagulation factor
13	76	38.4	622	1	TBHU	thrombin (EC 3.4.2
14	75	37.9	475	1	EXCH	coagulation factor
15	74	37.4	642	2	S53434	plasma protein S p
16	74	37.4	676	1	KXHUS	plasma protein S p
17	73	36.9	452	1	A30351	coagulation factor
18	73	36.9	459	2	JQ0419	coagulation factor
19	73	36.9	646	2	S38819	plasma protein S -
20	72	36.4	675	1	KXBOS	plasma protein S p
21	70	35.4	675	1	KXRTS	plasma protein S p
22	69	34.8	461	1	KFHU	coagulation factor
23	67	33.8	642	2	S53433	plasma protein S p
24	66	33.3	416	1	KFBO	coagulation factor
25	64	32.3	625	1	TBBO	thrombin (EC 3.4.2
26	63	31.8	675	1	KXMSS	plasma protein S p
27	61.5	31.1	396	1	KXBOZ	plasma protein Z -
28	57.5	29.0	422	1	KXHUZ	plasma protein Z p
29	56.5	28.5	594	2	D84859	probable MAP kinas
30	54.5	27.5	576	2	G96763	probable MAP kinas
31	54.5	27.5	603	2	C96575	probable MAP kinas
32	54	27.3	673	2	A48089	growth arrest-spec
33	53	26.8	674	2	I55476	growth potentiatin
34	53	26.8	730	2	G64062	primosomal replica
35	52	26.3	678	2	B48089	growth arrest-spec
36	51	25.8	1363	2	I58375	protein-tyrosine k
37	50	25.3	1298	2	A48999	protein-tyrosine k
38	50	25.3	1684	2	T02367	hypothetical prote
39	49	24.7	182	2	JC1189	tyrosine kinase re
40	49	24.7	367	2	B56598	endothelial kinase
41	49	24.7	1356	2	JC1402	protein-tyrosine k
42	49	24.7	1367	2	A41228	protein-tyrosine k
43	49	24.7	1379	2	JC4954	vascular endotheli
44	48.5	24.5	323	2	T25948	hypothetical prote
45	48.5	24.5	510	2	B82918	ammonium transport

ALIGNMENTS

RESULT 1
KXHU
protein C (activated) (EC 3.4.21.69) precursor - human
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A22331; A25426; A21781; A23789; A00927
R;Foster, D.C.; Yoshitake, S.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985
A;Title: The nucleotide sequence of the gene for human protein C.
A;Reference number: A22331; MUID:85270390; PMID:2991887
A;Accession: A22331
A;Molecule type: DNA
A;Residues: 1-461 <FOS1>
A;Cross-references: GB:M11228; NID:g190333; PIDN:AAA60166.1; PID:g190334
R;Plutzky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986
A;Title: Evolution and organization of the human protein C gene.
A;Reference number: A25426; MUID:86120978; PMID:3511471
A;Accession: A25426
A;Molecule type: DNA
A;Residues: 1-445,'L',446-461 <PLU>
A;Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332
R;Foster, D.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984
A;Title: Characterization of a cDNA coding for human protein C.
A;Reference number: A21781; MUID:84272714; PMID:6589623
A;Accession: A21781
A;Molecule type: mRNA
A;Residues: 'Q',107-461 <FOS2>
A;Cross-references: GB:K02059; NID:g190322; PIDN:AAA60164.1; PID:g190323
R;Beckmann, R.J.; Schmidt, R.J.; Santerre, R.F.; Plutzky, J.; Crabtree, G.R.; Long, G.L.
Nucleic Acids Res. 13, 5233-5247, 1985
A;Title: The structure and evolution of a 461 amino acid human protein C precursor and it
A;Reference number: A23789; MUID:85269639; PMID:2991859
A;Accession: A23789
A;Molecule type: mRNA
A;Residues: 1-461 <BEC>
A;Cross-references: GB:X02750; NID:g35689; PIDN:CAA26528.1; PID:g763120
R;Miletich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 265, 11397-11404, 1990
A;Title: Beta protein C is not glycosylated at asparagine 329. The rate of translation ma
A;Reference number: A44605; MUID:90293094; PMID:1694179
A;Contents: annotation; carbohydrate binding sites; activation peptide
A;Note: the alpha form of protein C is glycosylated at Asn-329, and the beta form is not
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of factor
A;Reference number: A44606; MUID:92184750; PMID:1544894
A;Contents: annotation; beta-hydroxyaspartic acid
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that ina
ivation of factor Va is strongly enhanced by complexing with protein S. Protein C also fa

```
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction,
C;Genetics:
A;Gene: GDB:PROC
A;Cross-references: GDB:120317; OMIM:176860
A;Map position: 2q13-2q21
A;Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-86/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-197/Product: protein C light chain #status predicted <LCH>
F;92-131/Domain: EGF homology <EG1>
F;140-175/Domain: EGF homology <EG2>
F;200-461/Product: protein C heavy chain #status predicted <HCH>
F;200-211/Domain: activation peptide #status experimental <APT>
F;212-445/Domain: trypsin homology <TRY>
F;48,49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp
F;59-64,92-105,101-120,122-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/D
F;106-111/Disulfide bonds: #status predicted
F;110/Binding site: carbohydrate (Thr) (covalent) #status absent
F;113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;139,290,355/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;211-212/Cleavage site: Arg-Leu (thrombin) #status experimental
F;253,299,402/Active site: His, Asp, Ser #status predicted
F;371/Binding site: carbohydrate (Asn) (covalent) (partial) #status atypical

  Query Match             86.9%;  Score 172;  DB 1;  Length 461;
  Best Local Similarity   75.0%;  Pred. No. 1e-19;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  | ||  ||||  || ||::|||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
JX0210
protein C (activated) (EC 3.4.21.69) precursor - mouse
N;Alternate names: vitamin K-dependent serine proteinase
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: JX0210
R;Tada, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.
J. Biochem. 111, 491-495, 1992
A;Title: Isolation and characterization of a mouse protein C cDNA.
A;Reference number: JX0210; MUID:92316897; PMID:1618739
A;Accession: JX0210
A;Molecule type: mRNA
A;Residues: 1-461 <TAD>
A;Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386
A;Experimental source: liver
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-33/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;34-41/Domain: propeptide #status predicted <PRO>
F;42-196,199-461/Product: protein C #status predicted <PRC>
F;42-196/Domain: light chain #status predicted <PCL>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;199-461/Domain: heavy chain #status predicted <PCH>
F;199-211/Domain: activation peptide #status predicted <ACT>
F;212-461/Product: vitamin K-dependent serine proteinase #status predicted <VIT>
F;212-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-319,238-254,373-387,398-426/Disulfide bonds: #stat
F;214,290,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;253,299,402/Active site: His, Asp, Ser #status predicted

  Query Match             67.2%;  Score 133;  DB 1;  Length 461;
  Best Local Similarity   56.8%;  Pred. No. 2e-13;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  |||| |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
S18994
protein C (activated) (EC 3.4.21.69) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C;Accession: S18994; S24312
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
submitted to the EMBL Data Library, February 1992
A;Description: The cDNA cloninig and mRNA expression of rat protein C.
A;Reference number: S18994
A;Accession: S18994
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
Biochim. Biophys. Acta 1131, 329-332, 1992
A;Title: The cDNA cloning and mRNA expression of rat protein C.
A;Reference number: S24312; MUID:92329550; PMID:1627650
A;Accession: S24312
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA2>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; glycoprotein; hydrolase; serine proteinase
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-461/Product: protein C #status predicted <PRC>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;213-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-320,239-255,373-387,398-426/Disulfide bonds: #stat
F;215,291,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;254,300,402/Active site: His, Asp, Ser #status predicted

  Query Match             66.7%;  Score 132;  DB 1;  Length 461;
  Best Local Similarity   56.8%;  Pred. No. 2.9e-13;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  |||| |: ||::|:|||||| |:
Db         42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
KXBO
protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Nov-1980 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A26250; A18385; A18386; A00928
R;Long, G.L.; Balagaje, R.M.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984
A;Title: Cloning and sequence of liver cDNA coding for bovine protein C.
A;Reference number: A26250; MUID:85014826; PMID:6091100
A;Accession: A26250
A;Molecule type: mRNA
A;Residues: 1-456 <LON>
R;Fernlund, P.; Stenflo, J.
```

```
J. Biol. Chem. 257, 12170-12179, 1982
A;Title: Amino acid sequence of the light chain of bovine protein C.
A;Reference number: A18385; MUID:83007325; PMID:6896876
A;Accession: A18385
A;Molecule type: protein
A;Residues: 40-194 <FER>
A;Note: 82-Lys was also found
R;Drakenberg, T.; Fernlund, P.; Roepstorff, P.; Stenflo, J.
Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983
A;Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.
A;Reference number: A19316; MUID:83169769; PMID:6572939
A;Contents: annotation; revision to residue 110
R;Stenflo, J.; Fernlund, P.
J. Biol. Chem. 257, 12180-12190, 1982
A;Title: Amino acid sequence of the heavy chain of bovine protein C.
A;Reference number: A18386; MUID:83007326; PMID:6896877
A;Accession: A18386
A;Molecule type: protein
A;Residues: 197-454,'PV' <STE>
R;Esmon, N.L.; DeBault, L.E.; Esmon, C.T.
J. Biol. Chem. 258, 5548-5553, 1983
A;Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless p
A;Reference number: A37541; MUID:83213513; PMID:6304092
A;Contents: annotation; activation; calcium binding
R;Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.
J. Biol. Chem. 258, 5554-5560, 1983
A;Title: Structural changes required for activation of protein C are induced by Ca2+ bin
A;Reference number: A37542; MUID:83213514; PMID:6406503
A;Contents: annotation; activation; calcium binding
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this reacti
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with stro
cognition of the thrombin-thrombomodulin complex.
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>
F;24-83/Domain: Gla domain homology <GLA>
F;30-39/Domain: propeptide #status predicted <PRO>
F;40-194/Product: protein C light chain #status experimental <LCH>
F;98-128/Domain: EGF homology <EG1>
F;137-172/Domain: EGF homology <EG2>
F;197-456/Product: protein C heavy chain #status experimental <HCH>
F;197-210/Domain: activation peptide #status experimental <APT>
F;211-440/Domain: trypsin homology <TRY>
F;45,46,53,55,58,59,62,64,65,68,74/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;119-128,137-148,144-157,159-172,180-318,237-253,368-382,393-421/Disulfide bonds: #stat
F;136,289,350/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;252,298,397/Active site: His, Asp, Ser #status predicted
F;366/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             58.1%;  Score 115;  DB 1;  Length 456;
  Best Local Similarity   47.6%;  Pred. No. 1.6e-10;
  Matches   20;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  ||  :: | |   :|:|  |: ||:: :||:|||
Db         40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 5
EXRT
coagulation factor Xa (EC 3.4.21.6) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Jan-1995 #sequence_revision 07-Feb-1997 #text_change 08-Dec-2000
C;Accession: S49075; JC4670; PS0191; PS0190; I62745
R;Stanton, C.; Ross, P.; Hutson, S.; Wallin, R.
Thromb. Res. 80, 63-73, 1995
A;Title: Evidence for competition between vitamin K-dependent clotting factors for intra
A;Reference number: A58498; MUID:96093366; PMID:8578539
A;Accession: S49075
A;Molecule type: mRNA
A;Residues: 1-482 <STA1>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Note: submitted to the EMBL Data Library, June 1994
A;Note: neither the complete nucleic acid sequence nor the complete translation are shown
R;Stanton, C.; Ross, R.P.; Hutson, S.; Wallin, R.
Gene 169, 269-273, 1996
A;Title: Processing and expression of rat and human clotting factor-X-encoding cDNAs.
A;Reference number: JC4670; MUID:96194815; PMID:8647460
A;Accession: JC4670
A;Molecule type: mRNA
A;Residues: 1-482 <STA2>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Experimental source: Cos-1 cell
R;Enjyoji, K.; Miyazaki, K.; Kato, H.
J. Biochem. 109, 890-898, 1991
A;Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat plas
A;Reference number: PS0190; MUID:92041742; PMID:1718949
A;Accession: PS0191
A;Molecule type: protein
A;Residues: 41-58,'X',60-65 <ENJ1>
A;Accession: PS0190
A;Molecule type: protein
A;Residues: 183-186,'X',188-207 <ENJ2>
R;Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A;Title: Analysis of the partial nucleotide sequences and deduced primary structures of t
A;Reference number: I46196; MUID:94222160; PMID:8168596
A;Accession: I62745
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 295-383,'G',385-455 <MUR>
A;Cross-references: GB:D21215; NID:g415309; PIDN:BAA04756.1; PID:g455396
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pre
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutami
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status predicted <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-482/Product: coagulation factor X heavy chain #status predicted <HCH>
F;183-231/Domain: activation peptide #status predicted <APT>
F;232-482/Product: coagulation factor Xa heavy chain #status predicted <ACT>
F;232-460/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #statu
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-340,238-243,259-275,388-402,413
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;187/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;208/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;231-232/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #statu
F;274,320,417/Active site: His, Asp, Ser #status predicted

  Query Match             54.0%;  Score 107;  DB 1;  Length 482;
  Best Local Similarity   40.9%;  Pred. No. 3.3e-09;
  Matches   18;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||   ::  :| | |:  || |  |: :||| : |  ||:|:
Db         41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 6
EXBO
coagulation factor Xa (EC 3.4.21.6) precursor - bovine
N;Alternate names: Stuart factor
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
```

C;Accession: A22867; A14997; A12030; A34412; S39414; A00925
R;Fung, M.R.; Campbell, R.M.; MacGillivray, T.A.
Nucleic Acids Res. 12, 4481-4492, 1984
A;Title: Blood coagulation factor X mRNA encodes a single polypeptide chain containing a
A;Reference number: A22867; MUID:84247315; PMID:6330671
A;Accession: A22867
A;Molecule type: mRNA
A;Residues: 1-487 <FUN>
A;Cross-references: GB:X00673; NID:g192; PIDN:CAA25286.1; PID:g193
R;Enfield, D.L.; Ericsson, L.H.; Fujikawa, K.; Walsh, K.A.; Neurath, H.; Titani, K.
Biochemistry 19, 659-667, 1980
A;Title: Amino acid sequence of the light chain of bovine factor X-1 (Stuart factor).
A;Reference number: A14997; MUID:80130563; PMID:6766735
A;Accession: A14997
A;Molecule type: protein
A;Residues: 41-102,'N',104-180 <ENF>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood cc
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Contents: annotation; revision to residue 103
R;Titani, K.; Fujikawa, K.; Enfield, D.L.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.
Proc. Natl. Acad. Sci. U.S.A. 72, 3082-3086, 1975
A;Title: Bovine factor X-1 (Stuart factor): amino-acid sequence of heavy chain.
A;Reference number: A12030; MUID:76053069; PMID:1059093
A;Accession: A12030
A;Molecule type: protein
A;Residues: 183-292,294-295,'GDE',299-334,336-348,'AE',351-354,356-441,'GKFG',446-492 <T
A;Note: carbohydrate binding sites and disulfide bonds were determined
R;Persson, E.; Selander, M.; Linse, S.; Drakenberg, T.; Oehlin, A.K.; Stenflo, J.
J. Biol. Chem. 264, 16897-16904, 1989
A;Title: Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal
A;Reference number: A34412; MUID:89380326; PMID:2789221
A;Accession: A34412
A;Molecule type: protein
A;Residues: 85-126 <PER>
A;Note: beta-hydroxyaspartic acid site
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39414
A;Molecule type: protein
A;Residues: 183-196;199-209;216-233 <INO>
A;Note: carbohydrate binding sites
R;Titani, K.; Hermodson, M.A.; Fujikawa, K.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.; D
Biochemistry 11, 4899-4903, 1972
A;Title: Bovine factor X-1a (activated Stuart factor). Evidence of homology with mammali
A;Reference number: A12453; MUID:73053314; PMID:4264286
A;Contents: annotation; active site
R;Fujikawa, K.; Titani, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 72, 3359-3363, 1975
A;Title: Activation of bovine factor X (Stuart factor): conversion of factor Xaalpha to
A;Reference number: A13504; MUID:76053121; PMID:1059122
A;Contents: annotation; activation
R;Sugo, T.; Bjork, I.; Holmgren, A.; Stenflo, J.
J. Biol. Chem. 259, 5705-5710, 1984
A;Title: Calcium-binding properties of bovine factor X lacking the gamma-carboxyglutamic
A;Reference number: A38024; MUID:84185716; PMID:6546930
A;Contents: annotation; calcium binding
R;Morita, T.; Jackson, C.M.
J. Biol. Chem. 261, 4008-4014, 1986
A;Reference number: A38025; MUID:86140210; PMID:3949800
A;Contents: annotation; sulfate binding
C;Comment: Factor Xa converts prothrombin to thrombin during blood clotting.
C;Comment: The two chains are formed from a single-chain precursor by the excision of tw
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway), c
activation.
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with strc
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
C;Genetics:
A;Gene: F10
A;Map position: 13q34

C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-180/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-492/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-233/Domain: activation peptide #status experimental <APT>
F;234-492/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;234-461/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-341/Disulfide bonds: #status p
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;200/Binding site: sulfate (Tyr) (covalent) (partial) #status experimental
F;208,485/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;233-234/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat
F;240-245,260-276,389-403,414-442/Disulfide bonds: #status experimental
F;275,321,418/Active site: His, Asp, Ser #status predicted

Query Match 52.0%; Score 103; DB 1; Length 492;
Best Local Similarity 40.9%; Pred. No. 1.5e-08;
Matches 18; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ::  :| | |:   |    |: :||| : |  ||||:
Db         41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84
```

RESULT 7
EXHU
coagulation factor Xa (EC 3.4.21.6) precursor [validated] - human
N;Alternate names: Stuart factor
C;Species: Homo sapiens (man)
C;Date: 15-Nov-1984 #sequence_revision 02-May-1994 #text_change 08-Dec-2000
C;Accession: A24478; JQ0917; A42485; A25853; A22208; A21284; A20362; S39415; I54051; A00
R;Leytus, S.P.; Foster, D.C.; Kurachi, K.; Davie, E.W.
Biochemistry 25, 5098-5102, 1986
A;Title: Gene for human Factor X: a blood coagulation factor whose gene organization is
A;Reference number: A24478; MUID:87026600; PMID:3768336
A;Accession: A24478
A;Molecule type: DNA
A;Residues: 1-488 <LEY>
A;Cross-references: GB:L29433; GB:M14327; NID:g459809; PIDN:AAA52764.1; PID:g182831
R;Messier, T.L.; Pittman, D.D.; Long, G.L.; Kaufman, R.J.; Church, W.R.
Gene 99, 291-294, 1991
A;Title: Cloning and expression in COS-1 cells of a full-length cDNA encoding human coagu
A;Reference number: JQ0917; MUID:91216473; PMID:1902434
A;Accession: JQ0917
A;Molecule type: mRNA
A;Residues: 1-488 <MES>
A;Cross-references: GB:M57285; NID:g182389; PIDN:AAA52421.1; PID:g182390
R;Miao, C.H.; Leytus, S.P.; Chung, D.W.; Davie, E.W.
J. Biol. Chem. 267, 7395-7401, 1992
A;Title: Liver-specific expression of the gene coding for human factor X, a blood coagul
A;Reference number: A42485; MUID:92218390; PMID:1313796
A;Accession: A42485
A;Molecule type: DNA
A;Residues: 1-15 <MIA>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:93780, NCBIP:93787)
R;Kaul, R.K.; Hildebrand, B.; Roberts, S.; Jagadeeswaran, P.
Gene 41, 311-314, 1986
A;Title: Isolation and characterization of human blood-coagulation factor X cDNA.
A;Reference number: A25853; MUID:86221713; PMID:3011603
A;Accession: A25853
A;Molecule type: mRNA
A;Residues: 19-284,'E',289-488 <KAU>

A;Cross-references: GB:M22613; NID:g180335; PIDN:AAA51984.1; PID:g180336
R;Fung, M.R.; Hay, C.W.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 3591-3595, 1985
A;Title: Characterization of an almost full-length cDNA coding for human blood coagulati
A;Reference number: A22208; MUID:85216545; PMID:2582420
A;Accession: A22208
A;Molecule type: mRNA
A;Residues: 13-441,'S',443-488 <FUN>
A;Cross-references: GB:K03194; NID:g182840; PIDN:AAA52490.1; PID:g182841
R;Leytus, S.P.; Chung, D.W.; Kisiel, W.; Kurachi, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 3699-3702, 1984
A;Title: Characterization of a cDNA coding for human factor X.
A;Reference number: A21284; MUID:84222026; PMID:6587384
A;Accession: A21284
A;Molecule type: mRNA
A;Residues: 13-284,'E',289-488 <LE2>
A;Cross-references: GB:K01886
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.; Sasagawa, T.; Howald, W.N.; Kwa, E.Y.; Weins
Biochemistry 22, 2875-2884, 1983
A;Title: Complete amino acid sequence of the light chain of human blood coagulation fact
A;Reference number: A20362; MUID:83257207; PMID:6871167
A;Accession: A20362
A;Molecule type: protein
A;Residues: 41-179 <MCM>
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39415
A;Molecule type: protein
A;Residues: 183-234 <INO>
A;Note: glycosylation sites
A;Note: identification and characterization of beta-hydroxyaspartic acid
R;Jagadeeswaran, P.; Reddy, S.V.; Rao, K.J.; Hamsabhushanam, K.; Lyman, G.
Gene 84, 517-519, 1989
A;Title: Cloning and characterization of the 5' end (exon 1) of the gene encoding human
A;Reference number: I54051; MUID:90128299; PMID:2612918
A;Accession: I54051
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-23 <RES>
A;Cross-references: GB:M33297; NID:g183860; PIDN:AAA52636.1; PID:g553330
R;Padmanabhan, K.; Padmanabhan, K.P.; Tulinsky, A.; Park, C.H.; Bode, W.; Huber, R.; Bla
J. Mol. Biol. 232, 947-966, 1993
A;Title: Structure of human des(1-45) factor Xa at 2.2 angstroms resolution.
A;Reference number: A49458; MUID:93360277; PMID:8355279
A;Contents: annotation; X-ray crystallography, 2.2 angstroms
C;Comment: The two chains held together by one disulfide bond are formed from a single-c
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway) or
C;Genetics:
A;Gene: GDB:F10
A;Cross-references: GDB:119890; OMIM:227600
A;Map position: 13q34-13q34
A;Introns: 24/1; 77/3; 86/1; 124/1; 150/3; 249/3; 289/1
A;Note: deficiency of this factor causes Stuart disease
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-488/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-234/Domain: activation peptide #status experimental <APT>
F;235-488/Product: coagulation factor Xa heavy chain #status experimental <ACT>
F;235-462/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;57-62/Disulfide bonds: #status predicted
F;90-101,95-110,112-121,129-140,136-149,151-164,172-342,241-246,261-277,390-404,415-443/
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;199,211/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;221,231/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;234-235/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #statu
F;276,322,419/Active site: His, Asp, Ser #status experimental

Query Match 51.5%; Score 102; DB 1; Length 488;
Best Local Similarity 40.9%; Pred. No. 2.1e-08;
Matches 18; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ::   | | |:   | :  |: :||| | |  ||:|:
Db         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84
```

RESULT 8
I46932
coagulation factor VII - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 12-Feb-1999
C;Accession: I46932
R;Brothers, A.B.; Clarke, B.J.; Sheffield, W.P.; Blajchman, M.A.
Thromb. Res. 69, 231-238, 1993
A;Title: Complete nucleotide sequence of the cDNA encoding rabbit coagulation factor VII.
A;Reference number: I46932; MUID:93190306; PMID:8383365
A;Accession: I46932
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-443 <BRO>
A;Cross-references: GB:S56300; NID:g266294; PID:g266295
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
F;24-83/Domain: Gla domain homology <GLA>
F;89-120/Domain: EGF homology <EG1>
F;130-166/Domain: EGF homology <EG2>
F;192-425/Domain: trypsin homology <TRY>

Query Match 47.5%; Score 94; DB 2; Length 443;
Best Local Similarity 43.9%; Pred. No. 3.7e-07;
Matches 18; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              |||||  ||  || | |   :| |  |: :|:  :  |  ||
Db         40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80
```

RESULT 9
KFHU7
coagulation factor VIIa (EC 3.4.21.21) precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 19-May-1994 #text_change 08-Dec-2000
C;Accession: A28322; A23819; A31186; B31186; S63524
R;O'Hara, P.J.; Grant, F.J.; Haldeman, B.A.; Gray, C.L.; Insley, M.Y.; Hagen, F.S.; Murra
Proc. Natl. Acad. Sci. U.S.A. 84, 5158-5162, 1987
A;Title: Nuclectide sequence of the gene coding for human factor VII, a vitamin K-depende
A;Reference number: A28322; MUID:87260948; PMID:3037537
A;Accession: A28322
A;Molecule type: DNA
A;Residues: 1-466 <OHA>
A;Cross-references: GB:J02933; NID:g180333; PIDN:AAA51983.1; PID:g180334
R;Hagen, F.S.; Gray, C.L.; O'Hara, P.; Grant, F.J.; Saari, G.C.; Woodbury, R.G.; Hart, C
Proc. Natl. Acad. Sci. U.S.A. 83, 2412-2416, 1986
A;Title: Characterization of a cDNA coding for human factor VII.
A;Reference number: A23819; MUID:86205965; PMID:3486420
A;Accession: A23819
A;Molecule type: mRNA
A;Residues: 1-466 <HAG>
A;Cross-references: GB:M13232; NID:g182799; PIDN:AAA88040.1; PID:g182801
R;Thim, L.; Bjoern, S.; Christensen, M.; Nicolaisen, E.M.; Lund-Hansen, T.; Pedersen, A.H
Biochemistry 27, 7785-7793, 1988
A;Title: Amino acid sequence and posttranslational modifications of human factor VII-a fr
A;Reference number: A90539; MUID:89088153; PMID:3264725
A;Accession: A31186
A;Molecule type: protein

```
A;Residues: 61-212 <THI>
A;Accession: B31186
A;Molecule type: protein
A;Residues: 213-466 <TH2>
R;Bjoern, S.; Foster, D.C.; Thim, L.; Wiberg, F.C.; Christensen, M.; Komiyama, Y.; Peder
J. Biol. Chem. 266, 11051-11057, 1991
A;Title: Human plasma and recombinant factor VII. Characterization of O-glycosylations a
A;Reference number: A40529; MUID:91250411; PMID:1904059
A;Contents: annotation; carbohydrate binding sites
R;Persson, E.; Petersen, L.C.
Eur. J. Biochem. 234, 293-300, 1995
A;Title: Structurally and functionally distinct Ca(2+) binding sites in the gamma-carbox
A;Reference number: S63524; MUID:96096752; PMID:8529655
A;Accession: S63524
A;Molecule type: protein
A;Residues: 61-65;99-103;105-109;213-217;308-312 <PER>
C;Genetics:
A;Gene: GDB:F7
A;Cross-references: GDB:119897; OMIM:227500
A;Map position: 13q34-13q34
A;Introns: 22/1; 44/1; 97/3; 106/1; 144/1; 191/1; 227/3; 269/1
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the prese
oagulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-60/Domain: propeptide #status predicted <PRO>
F;45-104/Domain: Gla domain homology <GLA>
F;61-212/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;110-141/Domain: EGF homology <EG1>
F;151-187/Domain: EGF homology <EG2>
F;213-466/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;213-447/Domain: trypsin homology <TRY>
F;66,67,74,76,79,80,85,86,89,95/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;77-82,110-121,115-130,132-141,151-162,158-172,174-187,195-322,219-224,238-254,370-389,
F;112,120/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;123/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status absent
F;205,382/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;212-213/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;253,302,404/Active site: His, Asp, Ser #status predicted
F;350-351/Cleavage site: Arg-Gly (coagulation factor Xa) #status predicted

  Query Match             46.5%;  Score 92;  DB 1;  Length 466;
  Best Local Similarity   46.3%;  Pred. No. 8.3e-07;
  Matches   19;  Conservative    4;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||:||  ||  || | |    | |  |: ||:| : |  ||
Db         61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 10
S10511
thrombin (EC 3.4.21.5) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 03-May-2002
C;Accession: S10511; A60576; B42696
R;Dihanich, M.; Monard, D.
Nucleic Acids Res. 18, 4251, 1990
A;Title: cDNA sequence of rat prothrombin.
A;Reference number: S10511; MUID:90332426; PMID:2377469
A;Accession: S10511
A;Molecule type: mRNA
A;Residues: 1-617 <DIH>
A;Cross-references: EMBL:X52835; NID:g56969; PIDN:CAA37017.1; PID:g56970
R;Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.
Endocrinology 126, 167-175, 1990
A;Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
A;Reference number: A60576; MUID:90091942; PMID:2293980
A;Accession: A60576
A;Molecule type: protein
A;Residues: 44-58 <HEN>
A;Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat ute
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: B42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 383-617,'E' <BAN>
A;Cross-references: GB:M81397
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-617/Product: prothrombin #status experimental <PMAT>
F;109-187/Domain: kringle homology <KR1>
F;215-292/Domain: kringle homology <KR2>
F;360-609/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,5
F;402,458,564/Active site: His, Asp, Ser #status predicted

  Query Match             39.6%;  Score 78.5;  DB 2;  Length 617;
  Best Local Similarity   40.0%;  Pred. No. 0.00017;
  Matches   18;  Conservative    5;  Mismatches   21;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||  :| | |:   | :  |    |   ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 11
A35827
thrombin (EC 3.4.21.5) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
C;Accession: A35827; A42696; S12081
R;Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.
DNA Cell Biol. 9, 487-498, 1990
A;Title: Characterization of the cDNA coding for mouse prothrombin and localization of t
A;Reference number: A35827; MUID:91025551; PMID:2222810
A;Accession: A35827
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-618 <DEG>
A;Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814
A;Experimental source: strain C57BL/6
A;Note: the data were obtained from females resulting from the cross of M. domesticus an
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: A42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 384-618,'E' <BAN>
A;Cross-references: GB:M81394
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-618/Product: prothrombin B #status predicted <MAT>
F;109-187/Domain: kringle homology <KR1>
F;215-293/Domain: kringle homology <KR2>
F;361-610/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,5
F;403,459,565/Active site: His, Asp, Ser #status predicted

  Query Match             39.6%;  Score 78.5;  DB 2;  Length 618;
```

```
  Best Local Similarity   40.0%;  Pred. No. 0.00017;
  Matches   18;  Conservative    5;  Mismatches   21;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||  :| | |:   | :  |     |   ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 12
KFBO7
coagulation factor VIIa (EC 3.4.21.21) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-May-1990 #sequence_revision 23-Mar-1995 #text_change 16-Jul-1999
C;Accession: A31979; C20274
R;Takeya, H.; Kawabata, S.; Nakagawa, K.; Yamamichi, Y.; Miyata, T.; Iwanaga, S.
J. Biol. Chem. 263, 14868-14877, 1988
A;Title: Bovine factor VII. Its purification and complete amino acid sequence.
A;Reference number: A31979; MUID:89008362; PMID:3049594
A;Accession: A31979
A;Molecule type: protein
A;Residues: 1-407 <TAK>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Accession: C20274
A;Molecule type: protein
A;Residues: 58-62,'X',64-68 <MCM>
A;Note: the residue designated 'X' was determined to be hydroxyaspartic acid
R;Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanaga,
J. Biochem. 104, 867-868, 1988
A;Title: A new trisaccharide sugar chain linked to a serine residue in bovine blood coag
A;Reference number: A44556; MUID:89213999; PMID:3149637
A;Contents: annotation
A;Note: structure and location of covalently bound carbohydrate
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the prese
gulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-152/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;1-44/Domain: Gla domain homology (fragment) <GLA>
F;50-81/Domain: EGF homology <EG1>
F;91-127/Domain: EGF homology <EG2>
F;153-407/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;153-387/Domain: trypsin homology <TRY>
F;6,7,14,16,19,20,25,26,29,34,35/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;17-22,50-61,55-70,72-81,91-102,98-112,114-127,135-262,159-164,178-194,310-329,340-368/
F;52/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;63/Modified site: erythro-beta-hydroxyaspartic acid (Asp) (partial) #status experiment
F;145,203/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;152-153/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;193,242,344/Active site: His, Asp, Ser #status predicted
F;290-291/Cleavage site: Arg-Gly (coagulation factor Xa) #status experimental

  Query Match             39.4%;  Score 78;  DB 1;  Length 407;
  Best Local Similarity   41.5%;  Pred. No. 0.00013;
  Matches   17;  Conservative    3;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  |  || | |   :| |  |  || : : |  ||
Db          1 ANGFLEELLPGSLERECREELCSFEEAHEIFRNEERTRQFW 41


RESULT 13
TBHU
thrombin (EC 3.4.21.5) precursor [validated] - human
N;Alternate names: coagulation factor II
N;Contains: prothrombin
C;Species: Homo sapiens (man)
C;Date: 30-Nov-1980 #sequence_revision 22-Jul-1994 #text_change 08-Dec-2000
C;Accession: A29351; A00914; B00914; A37549; A37550; I51952
R;Degen, S.J.F.; Davie, E.W.
Biochemistry 26, 6165-6177, 1987
A;Title: Nucleotide sequence of the gene for human prothrombin.
A;Reference number: A29351; MUID:88077877; PMID:2825773
A;Accession: A29351
A;Molecule type: DNA
A;Residues: 1-622 <DEG>
A;Cross-references: GB:M17262; GB:M33691; NID:g558069; PIDN:AAC63054.1; PID:g339641
R;Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.
Biochemistry 22, 2087-2097, 1983
A;Title: Characterization of the complementary deoxyribonucleic acid and gene coding for
A;Reference number: A00914; MUID:83231469; PMID:6305407
A;Accession: A00914
A;Molecule type: mRNA
A;Residues: 8-163,'N',165-622 <DE2>
A;Cross-references: GB:V00595; GB:J00307; NID:g37128; PIDN:CAA23842.1; PID:g1335344
A;Accession: B00914
A;Molecule type: DNA
A;Residues: 188-311 <DE3>
R;Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.
Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977
A;Reference number: A37549; MUID:77193964; PMID:266717
A;Accession: A37549
A;Molecule type: protein
A;Residues: 44-118,'N',120,'S',122-163,'I',165-175,'A',177-182,'T',184-193,'MV',196-308,'
R;Butkowski, R.J.; Elion, J.; Downing, M.R.; Mann, K.G.
J. Biol. Chem. 252, 4942-4957, 1977
A;Title: Primary structure of human prethrombin 2 and alpha-thrombin.
A;Reference number: A37550; MUID:77207112; PMID:873923
A;Accession: A37550
A;Molecule type: protein
A;Residues: 315-334,'N',336-348,'N',350-368,'N',370-397,'N',399-413,'N',415-484,'N',486-4
R;Rabiet, M.J.; Blashill, A.; Furie, B.; Furie, B.C.
J. Biol. Chem. 261, 13210-13215, 1986
A;Reference number: A37551; MUID:87008532; PMID:3759958
A;Contents: annotation; activation cleavages
R;MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.
Ann. N. Y. Acad. Sci. 485, 73-79, 1986
A;Title: Recombinant genetic approaches to functional mapping of thrombin.
A;Reference number: I51952; MUID:87182874; PMID:3471151
A;Accession: I51952
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2,'RI',5-100 <RES>
A;Cross-references: GB:M33031; NID:g190723; PIDN:AAA60220.1; PID:g190724
C;Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibri
C;Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds
) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chai
ter 314-Arg, are released in natural blood clotting.
C;Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.
C;Commemt: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxy
ent interaction with the negatively charged phospholipid membrane surface.
C;Comment: The prothrombin precursor is synthesized in the liver.
C;Genetics:
A;Gene: GDB:F2
A;Cross-references: GDB:119894; OMIM:176930
A;Map position: 11p11-11q12
A;Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552/
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; duplic
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-87/Domain: Gla domain homology <GLA>
F;44-622/Product: prothrombin #status experimental <MAT>
F;44-327/Domain: activation peptide #status experimental <APT>
F;108-186/Domain: kringle homology <KR1>
F;213-291/Domain: kringle homology <KR2>
F;328-363/Product: thrombin light chain #status experimental <LCH>
F;364-622/Product: thrombin heavy chain #status experimental <HCH>
F;364-613/Domain: trypsin homology <TRY>
F;49,50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #status e
F;60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds: #status p
F;121,143/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
F;336-482,536-550,564-594/Disulfide bonds: #status predicted
F;391-407/Disulfide bonds: #status experimental
F;406,462/Active site: His, Asp #status predicted
F;416/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;568/Active site: Ser #status experimental

  Query Match             38.4%;  Score 76;  DB 1;  Length 622;
  Best Local Similarity   34.1%;  Pred. No. 0.00042;
  Matches   15;  Conservative    7;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||:||  :|  :| | |:   | :  |    |    |  ||:|:
Db         44 ANTFLEEVRKGNLERECVEETCSYEEAFEALESSTATDVFWAKY 87


RESULT 14
EXCH
coagulation factor Xa (EC 3.4.21.6) precursor - chicken
N;Alternate names: virus-activating proteinase
C;Species: Gallus gallus (chicken)
C;Date: 12-Feb-1993 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C;Accession: S15838; S20380; S20381
R;Suzuki, H.; Harada, A.; Hayashi, Y.; Wada, K.; Asaka, J.; Gotoh, B.; Ogasawara, T.; Na
FEBS Lett. 283, 281-285, 1991
A;Title: Primary structure of the virus activating protease from chick embryo. Its ident
A;Reference number: S15838; MUID:91257322; PMID:2044767
A;Accession: S15838
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-475 <SUZ>
A;Cross-references: DDBJ:D00844; NID:g222869; PIDN:BAA00724.1; PID:g222870
R;Gotoh, B.; Yamauchi, F.; Ogasawara, T.; Nagai, Y.
FEBS Lett. 296, 274-278, 1992
A;Title: Isolation of factor Xa from chick embryo as the amniotic endoprotease responsib
A;Reference number: S20380; MUID:92164779; PMID:1537403
A;Accession: S20380
A;Molecule type: protein
A;Residues: 41-55 <GO2>
A;Accession: S20381
A;Molecule type: protein
A;Residues: 241-246,'X',248-251,'X',253-261 <GOT>
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-185/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-167/Domain: EGF homology <EG2>
F;186-475/Product: coagulation factor X heavy chain #status predicted <HCH>
F;186-240/Domain: activation peptide #status predicted <APT>
F;241-475/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;241-468/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-152,154-167,175-348,247-252,267-283,396-410,42
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;196,207,228,285/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;282,328,425/Active site: His, Asp, Ser #status predicted

  Query Match             37.9%;  Score 75;  DB 1;  Length 475;
  Best Local Similarity   35.7%;  Pred. No. 0.00046;
  Matches   15;  Conservative    7;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  ::  :: | |    |    |:  ||| : |  ||:
Db         41 ANSFLEEMKQGNIERECNEERCSKEEAREAFEDNEKTEEFWN 82


RESULT 15
S53434
plasma protein S precursor, vitamin K dependent - rhesus macaque (fragment)
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 16-Jul-1999
C;Accession: S53434
R;Greengard, J.S.; Fernandez, J.A.; Radtke, K.P.; Griffin, J.H.
Biochem. J. 305, 397-403, 1995
A;Title: Identification of candidate residues for interaction of protein S with C4b bind
A;Reference number: S53433; MUID:95134217; PMID:7832752
A;Accession: S53434
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-642 <GRE>
A;Cross-references: EMBL:L31380
A;Experimental source: tissue type liver
A;Note: the source is designated as rhesus monkey
C;Genetics:
A;Gene: PROS
C;Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat hom
F;1-51/Domain: Gla domain homology (fragment) <GLA>
F;1-7/Domain: signal sequence (fragment) #status predicted <SIG>
F;8-642/Product: plasma protein S #status predicted <MAT>
F;87-120/Domain: EGF homology <EG1>
F;127-165/Domain: EGF homology <EG2>
F;171-207/Domain: EGF homology <EG3>
F;213-248/Domain: EGF homology <EG4>
F;281-633/Domain: sex hormone-binding globulin homology <SHB>
F;291-444/Domain: laminin G repeat homology <LGR>

  Query Match             37.4%;  Score 74;  DB 2;  Length 642;
  Best Local Similarity   34.1%;  Pred. No. 0.00091;
  Matches   15;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| |   :  :| | ||  :|:   |: :||:  :|  |: |:
Db          8 ANSMLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 51


Search completed: March  1, 2004, 10:10:00
Job time : 14.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:58:27 ; Search time 10.5 Seconds
(without alignments)
218.199 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	172	86.9	461	1	PRTC_HUMAN	P04070 homo sapien
2	141	71.2	458	1	PRTC_RABIT	Q28661 oryctolagus
3	133	67.2	461	1	PRTC_MOUSE	P33587 mus musculu
4	132	66.7	461	1	PRTC_RAT	P31394 rattus norv
5	126	63.6	459	1	PRTC_PIG	Q9glp2 sus scrofa
6	115	58.1	456	1	PRTC_BOVIN	P00745 bos taurus
7	103	52.0	492	1	FA10_BOVIN	P00743 bos taurus
8	102	51.5	488	1	FA10_HUMAN	P00742 homo sapien
9	96	48.5	231	1	TMG3_HUMAN	Q9bzd7 homo sapien
10	95	48.0	490	1	FA10_RABIT	O19045 oryctolagus
11	94	47.5	444	1	FA7_RABIT	P98139 oryctolagus
12	92	46.5	466	1	FA7_HUMAN	P08709 homo sapien
13	81	40.9	218	1	TMG1_HUMAN	O14668 homo sapien
14	78.5	39.6	617	1	THRB_RAT	P18292 rattus norv
15	78.5	39.6	618	1	THRB_MOUSE	P19221 mus musculu
16	78	39.4	226	1	TMG4_HUMAN	Q9bzd6 homo sapien
17	78	39.4	407	1	FA7_BOVIN	P22457 bos taurus
18	77	38.9	376	1	FA10_HOPST	P83370 hoplocephal
19	76	38.4	622	1	THRB_HUMAN	P00734 homo sapien
20	75	37.9	376	1	FA10_TROCA	P81428 tropidechis
21	75	37.9	475	1	FA10_CHICK	P25155 gallus gall
22	74	37.4	649	1	PRTS_MACMU	Q28520 macaca mula
23	74	37.4	676	1	PRTS_HUMAN	P07225 homo sapien
24	73	36.9	446	1	FA7_MOUSE	P70375 mus musculu
25	73	36.9	452	1	FA9_CANFA	P19540 canis famil
26	73	36.9	459	1	FA9_MOUSE	P16294 mus musculu
27	73	36.9	646	1	PRTS_RABIT	P98118 oryctolagus
28	72	36.4	675	1	PRTS_BOVIN	P07224 bos taurus
29	70	35.4	675	1	PRTS_RAT	P53813 rattus norv
30	69	34.8	461	1	FA9_HUMAN	P00740 homo sapien
31	69	34.8	461	1	FA9_PANTR	Q95nd7 pan troglod
32	66	33.3	416	1	FA9_BOVIN	P00741 bos taurus
33	65	32.8	98	1	FA10_NOTSC	P82807 notechis sc
34	64	32.3	625	1	THRB_BOVIN	P00735 bos taurus
35	63	31.8	675	1	PRTS_MOUSE	Q08761 mus musculu
36	61.5	31.1	396	1	PRTZ_BOVIN	P00744 bos taurus
37	58	29.3	202	1	TMG2_HUMAN	O14669 homo sapien
38	57.5	29.0	400	1	PRTZ_HUMAN	P22891 homo sapien
39	53	26.8	730	1	PRIA_HAEIN	P44647 haemophilus
40	51	25.8	1363	1	VGR3_MOUSE	P35917 mus musculu
41	50	25.3	501	1	MKC1_CANAL	P43068 candida alb
42	50	25.3	1298	1	VGR3_HUMAN	P35916 homo sapien
43	49.5	25.0	941	1	VP11_HUMAN	Q9h270 homo sapien
44	49.5	25.0	941	1	VP11_MOUSE	Q91w86 mus musculu
45	49	24.7	1343	1	VGR2_RAT	O08775 rattus norv

ALIGNMENTS

RESULT 1
PRTC_HUMAN
ID PRTC_HUMAN STANDARD; PRT; 461 AA.
AC P04070; Q15189; Q15190; Q16001;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE factor XIV).
GN PROC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270390; PubMed=2991887;
RA Foster D.C., Yoshitake S., Davie E.W.;
RT "The nucleotide sequence of the gene for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85269639; PubMed=2991859;
RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutzky J., Crabtree G.R.,
RA Long G.L.;
RT "The structure and evolution of a 461 amino acid human protein C
RT precursor and its messenger RNA, based upon the DNA sequence of
RT cloned human liver cDNAs.";
RL Nucleic Acids Res. 13:5233-5247(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86120978; PubMed=3511471;
RA Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;
RT "Evolution and organization of the human protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 106-461 FROM N.A.
RX MEDLINE=84272714; PubMed=6589623;
RA Foster D.C., Davie E.W.;
RT "Characterization of a cDNA coding for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-371.
RX MEDLINE=90293094; PubMed=1694179;
RA Miletich J.P., Broze G.J. Jr.;
RT "Beta protein C is not glycosylated at asparagine 329. The rate of
RT translation may influence the frequency of usage at asparagine-X-
RT cysteine sites.";
RL J. Biol. Chem. 265:11397-11404(1990).
RN [8]
RP HYDROXYLATION.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [9]
RP 3D-STRUCTURE MODELING OF 175-450.
RX MEDLINE=94272342; PubMed=8003977;
RA Fisher C.L., Greengard J.S., Griffin J.H.;
RT "Models of the serine protease domain of the human antithrombotic
RT plasma factor activated protein C and its zymogen.";
RL Protein Sci. 3:588-599(1994).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.
RX MEDLINE=97157472; PubMed=9003757;
RA Mather T., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,
RA Bode W.;
RT "The 2.8 A crystal structure of Gla-domainless activated protein C.";
RL EMBO J. 15:6822-6831(1996).
RN [11]
RP REVIEW ON PROC VARIANTS.
RX MEDLINE=93190290; PubMed=8446940;
RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,
RA Sala N., Cooper D.N.;
RT "Protein C deficiency: a database of mutations. For the Protein C & S
RT Subcommittee of the Scientific and Standardization Committee of the
RT International Society on Thrombosis and Haemostasis.";
RL Thromb. Haemost. 69:77-84(1993).
RN [12]
RP VARIANT PROC DEFICIENCY CYS-444.
RX MEDLINE=87204221; PubMed=2437584;
RA Romeo G., Hassan H.J., Staempfli S., Roncuzzi L., Cianetti L.,
RA Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,
RA Cortese R.;
RT "Hereditary thrombophilia: identification of nonsense and missense
RT mutations in the protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).
RN [13]
RP VARIANT PROC DEFICIENCY TRP-211.
RX MEDLINE=90098906; PubMed=2602169;
RA Grundy C.B., Chitolie A., Talbot S., Bevan D., Kakkar V.V.,
RA Cooper D.N.;
RT "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in
RT the protein C gene causing thrombosis.";
RL Nucleic Acids Res. 17:10513-10513(1989).
RN [14]
RP VARIANT PROC DEFICIENCY CYS-272.
RX MEDLINE=91329836; PubMed=1868249;
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;
RT "The spectrum of genetic defects in a panel of 40 Dutch families with
RT symptomatic protein C deficiency type I: heterogeneity and founder
RT effects.";
RL Blood 78:890-894(1991).
RN [15]
RP VARIANTS PROC DEFICIENCY ALA-62 AND MET-76.
RX MEDLINE=92190481; PubMed=1347706;
RA Bovill E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,
RA Rainville I.R., Long G.L.;
RT "Protein CVermont: symptomatic type II protein C deficiency
RT associated with two GLA domain mutations.";
RL Blood 79:1456-1465(1992).
RN [16]
RP VARIANT PROC DEFICIENCY ASP-418.
RX MEDLINE=92305321; PubMed=1611081;
RA Sugahara Y., Miura O., Yuen P., Aoki N.;
RT "Protein C deficiency Hong Kong 1 and 2: hereditary protein C
RT deficiency caused by two mutant alleles, a 5-nucleotide deletion and
RT a missense mutation.";
RL Blood 80:126-133(1992).
RN [17]
RP VARIANT PROC DEFICIENCY LEU-289.
RX MEDLINE=92380660; PubMed=1511988;
RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;
RT "A novel homozygous missense mutation in the protein C (PROC) gene
RT causing recurrent venous thrombosis.";
RL Hum. Genet. 89:683-684(1992).
RN [18]
RP VARIANTS PROC DEFICIENCY GLN-220 AND TRP-220.
RX MEDLINE=92380661; PubMed=1511989;
RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;
RT "Two different missense mutations at Arg 178 of the protein C (PROC)
RT gene causing recurrent venous thrombosis.";
RL Hum. Genet. 89:685-686(1992).
RN [19]
RP VARIANT PROC DEFICIENCY GLN-220.
RX MEDLINE=93250852; PubMed=1301959;
RA Gandrille S., Vidaud M., Aiach M., Alhenc-Gelas M., Fischer A.M.,
RA Gouault-Heilman M., Toulon P., Fiessinger J.N., Goossens M.;
RT "Two novel mutations responsible for hereditary type I protein C
RT deficiency: characterization by denaturing gradient gel
RT electrophoresis.";
RL Hum. Mutat. 1:491-500(1992).
RN [20]
RP VARIANT PROC DEFICIENCY SER-334.
RX MEDLINE=92276939; PubMed=1593215;
RA Yamamoto K., Matsushita T., Sugiura I., Takamatsu J., Iwasaki E.,
RA Wada H., Deguchi K., Shirakawa S., Saito H.;
RT "Homozygous protein C deficiency: identification of a novel missense
RT mutation that causes impaired secretion of the mutant protein C.";
RL J. Lab. Clin. Med. 119:682-689(1992).
RN [21]
RP VARIANTS PROC DEFICIENCY TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.
RX MEDLINE=93313192; PubMed=8324221;
RA Gandrille S., Alhenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,
RA Juhan-Vague I., Aiach M.;
RT "Five novel mutations located in exons III and IX of the protein C
RT gene in patients presenting with defective protein C anticoagulant
RT activity.";
RL Blood 82:159-168(1993).
RN [22]
RP VARIANTS PROC DEFICIENCY GLY-14; GLN-211; TYR-244; GLN-253; LEU-321;
RP CYS-328; ILE-385; THR-388 AND VAL-388.
RX MEDLINE=93271391; PubMed=8499565;
RA Poort S.R., Pabinger-Fasching I., Mannhalter C., Reitsma P.H.,
RA Bertina R.M.;
RT "Twelve novel and two recurrent mutations in 14 Austrian families
RT with hereditary protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:273-280(1993).
RN [23]
RP VARIANT PROC DEFICIENCY TRP-57.
RX MEDLINE=93271396; PubMed=8499568;

```
  Query Match             86.9%;  Score 172;  DB 1;  Length 461;
  Best Local Similarity   75.0%;  Pred. No. 5.9e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  | ||  ||||  ||  ||::|||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
PRTC_RABIT
ID   PRTC_RABIT     STANDARD;      PRT;   458 AA.
AC   Q28661;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Shen L., He X., Dahlback B.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49933; AAA92956.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     27       BY SIMILARITY.
FT   PROPEP        28     36       BY SIMILARITY.
FT   CHAIN         37    458       VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN         37    192       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN        195    458       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE      195    209       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE         209    210       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN        91    126       EGF-LIKE 1.
FT   DOMAIN       130    170       EGF-LIKE 2.
FT   DOMAIN       210    458       SERINE PROTEASE.
FT   MOD_RES       42     42       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       43     43       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       50     50       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       52     52       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       56     56       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       62     62       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       65     65       GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES      107    107       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE     250    250       CHARGE RELAY SYSTEM.
FT   ACT_SITE     296    296       CHARGE RELAY SYSTEM.
FT   ACT_SITE     399    399       CHARGE RELAY SYSTEM.
FT   DISULFID      53     58       BY SIMILARITY.
FT   DISULFID      86    105       BY SIMILARITY.
FT   DISULFID      95    100       BY SIMILARITY.
FT   DISULFID      99    114       BY SIMILARITY.
FT   DISULFID     116    125       BY SIMILARITY.
FT   DISULFID     134    145       BY SIMILARITY.
FT   DISULFID     141    154       BY SIMILARITY.
FT   DISULFID     156    169       BY SIMILARITY.
FT   DISULFID     177    316       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     235    251       BY SIMILARITY.
FT   DISULFID     370    384       BY SIMILARITY.
FT   DISULFID     395    423       BY SIMILARITY.
FT   CARBOHYD     133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     287    287       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   458 AA;  51087 MW;  D75A5F990C8F29D7 CRC64;

  Query Match             71.2%;  Score 141;  DB 1;  Length 458;
  Best Local Similarity   61.4%;  Pred. No. 9e-16;
  Matches   27;  Conservative    4;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
                |||||  || ||| | |:  :||   || ||: |||||||| |:
Db        37 ANSFLEELRPSSLERECVEEVCDLEEAKEIFQSVDDTLAFWYKY 80


RESULT 3
PRTC_MOUSE
ID   PRTC_MOUSE     STANDARD;      PRT;   461 AA.
AC   P33587; O35498;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92316897; PubMed=1618739;
RA   Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;
RT   "Isolation and characterization of a mouse protein C cDNA.";
RL   J. Biochem. 111:491-495(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98152576; PubMed=9493582;
RA   Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,
RA   Castellino F.J.;
RT   "Nucleotide structure and characterization of the murine gene encoding
RT   anticoagulant protein C.";
RL   Thromb. Haemost. 79:310-316(1998).
RN   [3]
RP   SEQUENCE OF 274-434 FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=94318474; PubMed=8043441;
RA   Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT   "A comparative study of partial primary structures of the catalytic
RT   region of mammalian protein C.";
RL   Br. J. Haematol. 86:590-600(1994).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D10445; BAA01235.1; -.
DR   EMBL; AF034569; AAC33795.1; -.
DR   EMBL; D43755; BAA07812.1; -.
DR   PIR; JX0210; JX0210.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   MGD; MGI:97771; Proc.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL        1     33       BY SIMILARITY.
FT   PROPEP       34     41       BY SIMILARITY.
FT   CHAIN        42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN       199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE     199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE        212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN       96    131       EGF-LIKE 1.
FT   DOMAIN      135    175       EGF-LIKE 2.
FT   DOMAIN      213    461       SERINE PROTEASE.
FT   MOD_RES      47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES     112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE    253    253       CHARGE RELAY SYSTEM.
FT   ACT_SITE    299    299       CHARGE RELAY SYSTEM.
FT   ACT_SITE    402    402       CHARGE RELAY SYSTEM.
FT   DISULFID     58     63       BY SIMILARITY.
FT   DISULFID     91    110       BY SIMILARITY.
FT   DISULFID    100    105       BY SIMILARITY.
FT   DISULFID    104    119       BY SIMILARITY.
FT   DISULFID    121    130       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
FT   DISULFID    146    159       BY SIMILARITY.
FT   DISULFID    161    174       BY SIMILARITY.
```

```
FT   DISULFID    182    319       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    238    254       BY SIMILARITY.
FT   DISULFID    373    387       BY SIMILARITY.
FT   DISULFID    398    426       BY SIMILARITY.
FT   CARBOHYD    214    214       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    328    328       MISSING (IN REF. 2).
FT   CONFLICT    393    393       N -> D (IN REF. 2).
SQ   SEQUENCE   461 AA;  51945 MW;  53FAA0D85B194D6E CRC64;

  Query Match             67.2%;  Score 133;  DB 1;  Length 461;
  Best Local Similarity   56.8%;  Pred. No. 2e-14;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  ||||  |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
PRTC_RAT
ID   PRTC_RAT       STANDARD;      PRT;   461 AA.
AC   P31394;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=92329550; PubMed=1627650;
RA   Okafuji T., Maekawa K., Nawa K., Marumoto Y.;
RT   "The cDNA cloning and mRNA expression of rat protein C.";
RL   Biochim. Biophys. Acta 1131:329-332(1992).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X64336; CAA45617.1; -.
DR   PIR; S18994; S18994.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     32       BY SIMILARITY.
FT   PROPEP      33     41       BY SIMILARITY.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN      199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE    199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     213    461       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   254    254       CHARGE RELAY SYSTEM.
FT   ACT_SITE   300    300       CHARGE RELAY SYSTEM.
FT   ACT_SITE   402    402       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
FT   DISULFID   100    105       BY SIMILARITY.
FT   DISULFID   104    119       BY SIMILARITY.
FT   DISULFID   121    130       BY SIMILARITY.
FT   DISULFID   139    150       BY SIMILARITY.
FT   DISULFID   146    159       BY SIMILARITY.
FT   DISULFID   161    174       BY SIMILARITY.
FT   DISULFID   182    320       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   239    255       BY SIMILARITY.
FT   DISULFID   373    387       BY SIMILARITY.
FT   DISULFID   398    426       BY SIMILARITY.
```

```
FT   CARBOHYD    215    215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   461 AA;  51912 MW;  8A4CF93664EDACD5 CRC64;

  Query Match             66.7%;  Score 132;  DB 1;  Length 461;
  Best Local Similarity   56.8%;  Pred. No. 2.9e-14;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  ||||  |: ||::|:|||||| |:
Db         42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 5
PRTC_PIG
ID   PRTC_PIG       STANDARD;      PRT;   459 AA.
AC   Q9GLP2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21121490; PubMed=11229814;
RA   Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
RA   Kim H.K.W.;
RT   "Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT   protein modeling of membrane binding sites and comparative anatomy of
RT   domains.";
RL   Cell. Mol. Life Sci. 58:148-159(2001).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF191307; AAG28380.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     18       BY SIMILARITY.
FT   PROPEP      19     41       BY SIMILARITY.
FT   CHAIN       42    459       VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY
FT                               SIMILARITY).
FT   CHAIN      199    459       PROTEIN C HEAVY CHAIN (BY
FT                               SIMILARITY).
FT   PEPTIDE    199    213       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       213    214       CLEAVAGE (BY THROMBIN) (BY
FT                               SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     214    459       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   255    255       CHARGE RELAY SYSTEM.
FT   ACT_SITE   301    301       CHARGE RELAY SYSTEM.
FT   ACT_SITE   400    400       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
FT   DISULFID   100    105       BY SIMILARITY.
FT   DISULFID   104    119       BY SIMILARITY.
FT   DISULFID   121    130       BY SIMILARITY.
FT   DISULFID   139    150       BY SIMILARITY.
FT   DISULFID   146    159       BY SIMILARITY.
FT   DISULFID   161    174       BY SIMILARITY.
FT   DISULFID   182    321       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   240    256       BY SIMILARITY.
FT   DISULFID   371    385       BY SIMILARITY.
```

```
FT   DISULFID    396    424       BY SIMILARITY.
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   459 AA;  51866 MW;  8541AAC14CC16D09 CRC64;

  Query Match             63.6%;  Score 126;  DB 1;  Length 459;
  Best Local Similarity   54.5%;  Pred. No. 2.9e-13;
  Matches   24;  Conservative    7;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  || ||| | |    |||  |: ||:: ::|:|||||:
Db         42 ANSFLEELRPSSLERECKEETCDFEEAREIFQNTENTMAFWSKY 85


RESULT 6
PRTC_BOVIN
ID   PRTC_BOVIN     STANDARD;      PRT;   456 AA.
AC   P00745;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85014826; PubMed=6091100;
RA   Long G.L., Balagaje R.M., McGillivray R.T.A.;
RT   "Cloning and sequencing of liver cDNA coding for bovine protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).
RN   [2]
RP   SEQUENCE OF 40-194, AND CARBOHYDRATE-LINKAGE SITE ASN-136.
RX   MEDLINE=83007325; PubMed=6896876;
RA   Fernlund P., Stenflo J.;
RT   "Amino acid sequence of the light chain of bovine protein C.";
RL   J. Biol. Chem. 257:12170-12179(1982).
RN   [3]
RP   REVISION TO 110.
RX   MEDLINE=83169769; PubMed=6572939;
RA   Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;
RT   "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).
RN   [4]
RP   SEQUENCE OF 197-456, AND CARBOHYDRATE-LINKAGE SITES ASN-289; ASN-350
RP   AND ASN-366.
RX   MEDLINE=83007326; PubMed=6896877;
RA   Stenflo J., Fernlund P.;
RT   "Amino acid sequence of the heavy chain of bovine protein C.";
RL   J. Biol. Chem. 257:12180-12190(1982).
RN   [5]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213513; PubMed=6304092;
RA   Esmon N.L., Debault L.E., Esmon C.T.;
RT   "Proteolytic formation and properties of gamma-carboxyglutamic acid-
RT   domainless protein C.";
RL   J. Biol. Chem. 258:5548-5553(1983).
RN   [6]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213514; PubMed=6406503;
RA   Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;
RT   "Structural changes required for activation of protein C are induced
RT   by Ca2+ binding to a high affinity site that does not contain gamma-
RT   carboxyglutamic acid.";
RL   J. Biol. Chem. 258:5554-5560(1983).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K02435; AAA30685.1; -.
DR   PIR; A26250; KXBO.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     29
FT   PROPEP       30     39
FT   CHAIN        40    194       PROTEIN C LIGHT CHAIN.
FT   CHAIN       197    456       PROTEIN C HEAVY CHAIN.
FT   PEPTIDE     197    210       ACTIVATION PEPTIDE.
FT   DOMAIN       94    129       EGF-LIKE 1.
FT   DOMAIN      133    173       EGF-LIKE 2.
FT   DOMAIN      211    456       SERINE PROTEASE.
FT   MOD_RES      45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID.
```

```
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      62     62       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     110    110       HYDROXYLATION.
FT   ACT_SITE    252    252       CHARGE RELAY SYSTEM.
FT   ACT_SITE    298    298       CHARGE RELAY SYSTEM.
FT   ACT_SITE    397    397       CHARGE RELAY SYSTEM.
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    108       BY SIMILARITY.
FT   DISULFID     98    103       BY SIMILARITY.
FT   DISULFID    102    117       BY SIMILARITY.
FT   DISULFID    119    128       BY SIMILARITY.
FT   DISULFID    137    148       BY SIMILARITY.
FT   DISULFID    144    157       BY SIMILARITY.
FT   DISULFID    159    172       BY SIMILARITY.
FT   DISULFID    180    318       INTERCHAIN.
FT   DISULFID    237    253
FT   DISULFID    368    382
FT   DISULFID    393    421
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .).
FT   VARIANT      82     82       F -> K.
FT   CONFLICT    455    456       VP -> PV (IN REF. 4).
SQ   SEQUENCE   456 AA;  51407 MW;  CAAF6833F894C209 CRC64;

  Query Match             58.1%;  Score 115;  DB 1;  Length 456;
  Best Local Similarity   47.6%;  Pred. No. 2e-11;
  Matches   20;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  ||  :: | |   :|:|  |: ||:: :||:|||| 
Db         40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 7
FA10_BOVIN
ID   FA10_BOVIN     STANDARD;      PRT;   492 AA.
AC   P00743;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE OF 1-487 FROM N.A.
RX   MEDLINE=84247315; PubMed=6330671;
RA   Fung M.R., Campbell R.M., McGillivray R.T.A.;
RT   "Blood coagulation factor X mRNA encodes a single polypeptide chain
RT   containing a prepro leader sequence.";
RL   Nucleic Acids Res. 12:4481-4492(1984).
RN   [2]
RP   SEQUENCE OF 41-180.
RX   MEDLINE=80130563; PubMed=6766735;
RA   Enfield D.L., Ericsson L.H., Fujikawa K., Walsh K.A., Neurath H.,
RA   Titani K.;
RT   "Amino acid sequence of the light chain of bovine factor X1 (Stuart
RT   factor).";
RL   Biochemistry 19:659-667(1980).
RN   [3]
RP   REVISION TO 103.
RX   MEDLINE=83308813; PubMed=6688526;
RA   McMullen B.A., Fujikawa K., Kisiel W.;
RT   "The occurrence of beta-hydroxyaspartic acid in the vitamin
RT   K-dependent blood coagulation zymogens.";
RL   Biochem. Biophys. Res. Commun. 115:8-14(1983).
RN   [4]
RP   SEQUENCE OF 183-492, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX   MEDLINE=76053069; PubMed=1059093;
RA   Titani K., Fujikawa K., Enfield D.L., Ericsson L.H., Walsh K.A.,
RA   Neurath H.;
RT   "Bovine factor X1 (Stuart factor): amino-acid sequence of heavey
RT   chain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3082-3086(1975).
RN   [5]
RP   SEQUENCE OF 183-233, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=94062825; PubMed=8243461;
RA   Inoue K., Morita T.;
RT   "Identification of O-linked oligosaccharide chains in the activation
RT   peptides of blood coagulation factor X. The role of the carbohydrate
RT   moieties in the activation of factor X.";
RL   Eur. J. Biochem. 218:153-163(1993).
RN   [6]
RP   ACTIVE SITE.
RX   MEDLINE=73053314; PubMed=4264286;
RA   Titani K., Hermodson M.A., Fujikawa K., Ericsson L.H., Walsh K.A.,
RA   Neurath H., Davie E.W.;
RT   "Bovine factor X 1a (activated Stuart factor). Evidence of homology
RT   with mammalian serine proteases.";
RL   Biochemistry 11:4899-4903(1972).
RN   [7]
RP   PROCESSING.
RX   MEDLINE=76053121; PubMed=1059122;
RA   Fujikawa K., Titani K., Davie E.W.;
RT   "Activation of bovine factor X (Stuart factor): conversion of factor
RT   Xa-alpha to factor Xa-beta.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3359-3363(1975).
RN   [8]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=84185716; PubMed=6546930;
RA   Sugo T., Bjoerk I., Holmgren A., Stenflo J.;
RT   "Calcium-binding properties of bovine factor X lacking the gamma-
RT   carboxyglutamic acid-containing region.";
RL   J. Biol. Chem. 259:5705-5710(1984).
RN   [9]
RP   SULFATION.
RX   MEDLINE=86140210; PubMed=3949800;
RA   Morita T., Jackson C.M.;
RT   "Localization of the structural difference between bovine blood
RT   coagulation factors X1 and X2 to tyrosine 18 in the activation
RT   peptide.";
RL   J. Biol. Chem. 261:4008-4014(1986).
RN   [10]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=91084483; PubMed=2261466;
RA   Selander M., Persson E., Stenflo J., Drakenberg T.;
RT   "1H NMR assignment and secondary structure of the Ca2(+)-free form of
RT   the amino-terminal epidermal growth factor like domain in coagulation
RT   factor X.";
RL   Biochemistry 29:8111-8118(1990).
RN   [11]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92329412; PubMed=1627540;
RA   Ullner M., Selander M., Persson E., Stenflo J., Drakenberg T.,
RA   Teleman O.;
RT   "Three-dimensional structure of the apo form of the N-terminal
RT   EGF-like module of blood coagulation factor X as determined by NMR
RT   spectroscopy and simulated folding.";
RL   Biochemistry 31:5974-5983(1992).
RN   [12]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92406922; PubMed=1527084;
RA   Selander-Sunnerhagen M., Ullner M., Persson E., Teleman O.,
RA   Stenflo J., Drakenberg T.;
RT   "How an epidermal growth factor (EGF)-like domain binds calcium. High
RT   resolution NMR structure of the calcium form of the NH2-terminal EGF-
```

```
RT   like domain in coagulation factor X.";
RL   J. Biol. Chem. 267:19642-19649(1992).
RN   [13]
RP   STRUCTURE BY NMR OF 41-126.
RX   MEDLINE=96387194; PubMed=8794734;
RA   Sunnerhagen M., Olah G.A., Stenflo J., Forsen S., Drakenberg T.,
RA   Trewhella J.;
RT   "The relative orientation of Gla and EGF domains in coagulation
RT   factor X is altered by Ca2+ binding to the first EGF domain. A
RT   combined NMR-small angle X-ray scattering study.";
RL   Biochemistry 35:11547-11559(1996).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC       by the excision of two Arg residues and are held together by 1 or
CC       more disulfide bonds.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium.
CC   -!- PTM: N- and O-glycosylated.
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00673; CAA25286.1; -.
DR   PIR; A22867; EXBO.
DR   PDB; 1APO; 31-JAN-94.
DR   PDB; 1CCF; 31-MAY-94.
DR   PDB; 1WHE; 15-MAY-97.
DR   PDB; 1WHF; 15-MAY-97.
DR   PDB; 1IOD; 21-JAN-03.
DR   PDB; 1KIG; 28-OCT-98.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00743; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat; Sulfation; 3D-structure.
FT   SIGNAL       1     23       POTENTIAL.
FT   PROPEP      24     40
FT   CHAIN       41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN      183    492       FACTOR X HEAVY CHAIN.
FT   PROPEP     183    233       ACTIVATION PEPTIDE.
FT   CHAIN      234    492       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   PROPEP     476    492       MAY BE REMOVED BUT IS NOT NECESSARY FOR
FT                               ACTIVATION.
FT   DOMAIN      86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125    165       EGF-LIKE 2.
FT   DOMAIN     234    492       SERINE PROTEASE.
FT   ACT_SITE   275    275       CHARGE RELAY SYSTEM.
FT   ACT_SITE   321    321       CHARGE RELAY SYSTEM.
FT   ACT_SITE   418    418       CHARGE RELAY SYSTEM.
FT   MOD_RES     46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     54     54       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     56     56       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID.

  Query Match             52.0%;  Score 103;  DB 1;  Length 492;
  Best Local Similarity   40.9%;  Pred. No. 2.3e-09;
  Matches   18;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ::  :| | |:    |    |: :||| : |  ||||:
Db         41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84


RESULT 8
FA10_HUMAN
ID   FA10_HUMAN     STANDARD;      PRT;   488 AA.
AC   P00742; Q14340;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91216473; PubMed=1902434;
RA   Messier T.L., Pittman D.D., Long G.L., Kaufman R.J., Church W.R.;
RT   "Cloning and expression in COS-1 cells of a full-length cDNA encoding
RT   human coagulation factor X.";
RL   Gene 99:291-294(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87026600; PubMed=3768336;
RA   Leytus S.P., Foster D.C., Kurachi K., Davie E.W.;
RT   "Gene for human factor X: a blood coagulation factor whose gene
RT   organization is essentially identical with that of factor IX and
RT   protein C.";
RL   Biochemistry 25:5098-5102(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 13-488 FROM N.A.
RX MEDLINE=85216545; PubMed=2582420;
RA Fung M.R., Hay C.W., McGillivray R.T.A.;
RT "Characterization of an almost full-length cDNA coding for human
RT blood coagulation factor X.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3591-3595(1985).
RN [6]
RP SEQUENCE OF 19-488 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86221713; PubMed=3011603;
RA Kaul R.K., Hildebrand B., Roberts S., Jagadeeswaran P.;
RT "Isolation and characterization of human blood-coagulation factor X
RT cDNA.";
RL Gene 41:311-314(1986).
RN [7]
RP SEQUENCE OF 41-179.
RX MEDLINE=83257207; PubMed=6871167;
RA McMullen B.A., Fujikawa K., Kisiel W., Sasagawa T., Howald W.N.,
RA Kwa E.Y., Weinstein B.;
RT "Complete amino acid sequence of the light chain of human blood
RT coagulation factor X: evidence for identification of residue 63 as
RT beta-hydroxyaspartic acid.";
RL Biochemistry 22:2875-2884(1983).
RN [8]
RP SEQUENCE OF 115-488 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX MEDLINE=84222026; PubMed=6587384;
RA Leytus S.P., Chung D.W., Kisiel W., Kurachi K., Davie E.W.;
RT "Characterization of a cDNA coding for human factor X.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3699-3702(1984).
RN [9]
RP SEQUENCE OF 183-234, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=94062825; PubMed=8243461;
RA Inoue K., Morita T.;
RT "Identification of O-linked oligosaccharide chains in the activation
RT peptides of blood coagulation factor X. The role of the carbohydrate
RT moieties in the activation of factor X.";
RL Eur. J. Biochem. 218:153-163(1993).
RN [10]
RP SEQUENCE OF 1-23 FROM N.A.
RX MEDLINE=90128299; PubMed=2612918;
RA Jagadeeswaran P., Reddy S.V., Rao K.J., Hamsabhushanam K., Lyman G.;
RT "Cloning and characterization of the 5' end (exon 1) of the gene
RT encoding human factor X.";
RL Gene 84:517-519(1989).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-179 AND 235-278.
RX MEDLINE=93360277; PubMed=8355279;
RA Padmanabhan K., Padmanabhan K.P., Tulinsky A., Park C.H., Bode W.,
RA Huber R., Blankenship D.T., Cardin A.D., Kisiel W.;
RT "Structure of human des(1-45) factor Xa at 2.2-A resolution.";
RL J. Mol. Biol. 232:947-966(1993).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-179 AND 235-278.
RX MEDLINE=98283982; PubMed=9618463;
RA Kamata K., Kawamoto H., Honma T., Iwama T., Kim S.H.;
RT "Structural basis for chemical inhibition of human blood coagulation
RT factor Xa.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6630-6635(1998).
RN [13]
RP VARIANTS ILE-7 AND HIS-30.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [14]
RP ERRATUM.
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
CC -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC converts prothrombin to thrombin in the presence of factor Va,
CC calcium and phospholipid during blood clotting.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC Arg-|-Ile bonds in prothrombin to form thrombin.
CC -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC by the excision of two Arg residues and are held together by 1 or
CC more disulfide bonds.
CC -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC glutamate residues allows the modified protein to bind calcium.
CC -!- PTM: N- and O-glycosylated.
CC -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC --

CC --
DR EMBL; K03194; AAA52490.1; -.
DR EMBL; M57285; AAA52421.1; -.
DR EMBL; AF503510; AAM19347.1; -.
DR EMBL; BC046125; AAH46125.1; -.
DR EMBL; L29433; AAA52764.1; -.
DR EMBL; L00390; AAA52764.1; JOINED.
DR EMBL; L00391; AAA52764.1; JOINED.
DR EMBL; L00392; AAA52764.1; JOINED.
DR EMBL; L00393; AAA52764.1; JOINED.
DR EMBL; L00394; AAA52764.1; JOINED.
DR EMBL; L00395; AAA52764.1; JOINED.
DR EMBL; L00396; AAA52764.1; JOINED.
DR EMBL; M22613; AAA51984.1; -.
DR EMBL; K01886; AAA52486.1; -.
DR EMBL; M33297; AAA52636.1; -.
DR PIR; A24478; EXHU.
DR PDB; 1HCG; 08-MAY-95.
DR PDB; 1FAX; 29-OCT-97.
DR PDB; 1FXY; 17-JUN-98.
DR PDB; 1XKA; 23-MAR-99.
DR PDB; 1XKB; 23-MAR-99.
DR PDB; 1EZQ; 20-SEP-00.
DR PDB; 1F0R; 20-SEP-00.
DR PDB; 1F0S; 20-SEP-00.
DR PDB; 1FJS; 17-NOV-00.
DR PDB; 1G2L; 20-OCT-01.
DR PDB; 1G2M; 20-OCT-01.

```
DR   PDB; 1KSN; 19-JUN-02.
DR   PDB; 1KYE; 11-FEB-03.
DR   PDB; 1MQ5; 28-JAN-03.
DR   PDB; 1MQ6; 28-JAN-03.
DR   PDB; 1NFU; 25-FEB-03.
DR   PDB; 1NFW; 25-FEB-03.
DR   PDB; 1NFX; 25-FEB-03.
DR   PDB; 1NFY; 25-FEB-03.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00742; -.
DR   Genew; HGNC:3528; F10.
DR   MIM; 134530; -.
DR   MIM; 227600; -.
DR   GO; GO:0005576; C:extracellular; TAS.
DR   GO; GO:0003804; F:blood coagulation factor X activity; TAS.
DR   GO; GO:0007596; P:blood coagulation; TAS.

  Query Match             51.5%;  Score 102;  DB 1;  Length 488;
  Best Local Similarity   40.9%;  Pred. No. 3.3e-09;
  Matches   18;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||   ::    | | |:   | :  |: :||| | |  ||:|:
Db         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84


RESULT 9
TMG3_HUMAN
ID   TMG3_HUMAN     STANDARD;      PRT;   231 AA.
AC   Q9BZD7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 3 precursor.
GN   TMG3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spinal cord;
RX   MEDLINE=21117044; PubMed=11171957;
RA   Kulman J.D., Harris J.E., Xie L., Davie E.W.;
RT   "Identification of two novel transmembrane gamma-carboxyglutamic acid
RT   proteins expressed broadly in fetal and adult tissues.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in brain, lung, kidney and heart.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamate by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF326350; AAK00955.1; -.
DR   HSSP; P00740; 1CFH.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     19       POTENTIAL.
FT   CHAIN          20    231       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 3.
FT   DOMAIN         20     78       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       79    101       POTENTIAL.
FT   DOMAIN        102    231       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         23     60       GLA-RICH.
SQ   SEQUENCE   231 AA;  25848 MW;  8A373E4848490D81 CRC64;

  Query Match             48.5%;  Score 96;  DB 1;  Length 231;
  Best Local Similarity   39.0%;  Pred. No. 1.4e-08;
  Matches   16;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  ||  :: | |:  || :   | :||: : |: ||
Db         20 ANEFLEELRQGTIERECMEEICSYEEVKEVFENKEKTMEFW 60


RESULT 10
FA10_RABIT
ID   FA10_RABIT     STANDARD;      PRT;   490 AA.
AC   O19045;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97256311; PubMed=9101642;
RA   Pendurthi U.R., Anderson K.D., James H.L.;
RT   "Characterization of a full-length cDNA for rabbit factor X.";
RL   Thromb. Res. 85:503-514(1997).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC       by the excision of two Arg residues and are held together by 1 or
CC       more disulfide bonds.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium (By
CC       similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC       (BY SIMILARITY).
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003200; AAB62542.1; -.
DR   HSSP; P00742; 1HCG.
DR   MEROPS; S01.216; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
```

```
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat.
FT   SIGNAL       1     20       POTENTIAL.
FT   PROPEP      21     40       BY SIMILARITY.
FT   CHAIN       41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN      184    490       FACTOR X HEAVY CHAIN.
FT   PROPEP     184    232       ACTIVATION PEPTIDE.
FT   CHAIN      233    490       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   DOMAIN      86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125    165       EGF-LIKE 2.
FT   DOMAIN     233    490       SERINE PROTEASE.
FT   MOD_RES     46     46       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     54     54       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     56     56       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     59     59       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     65     65       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     69     69       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     72     72       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     75     75       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     79     79       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES    103    103       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   274    274       CHARGE RELAY SYSTEM.
FT   ACT_SITE   320    320       CHARGE RELAY SYSTEM.
FT   ACT_SITE   417    417       CHARGE RELAY SYSTEM.
FT   DISULFID    90    101       BY SIMILARITY.
FT   DISULFID    95    110       BY SIMILARITY.
FT   DISULFID   112    121       BY SIMILARITY.
FT   DISULFID   129    140       BY SIMILARITY.
FT   DISULFID   136    149       BY SIMILARITY.
FT   DISULFID   151    164       BY SIMILARITY.
FT   DISULFID   172    340       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   239    244       BY SIMILARITY.
FT   DISULFID   259    275       BY SIMILARITY.
FT   DISULFID   388    402       BY SIMILARITY.
FT   DISULFID   413    441       BY SIMILARITY.
FT   CARBOHYD    61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   490 AA;  53965 MW;  3A39FA85AF2A6D11 CRC64;

  Query Match             48.0%;  Score 95;  DB 1;  Length 490;
  Best Local Similarity   40.9%;  Pred. No. 4.9e-08;
  Matches   18;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |: :| | |:   | :  |  :||| : |  ||:|:
Db         41 ANSFLEELKKGNLERECMEENCSYEEALEVFEDREKTNEFWNKY 84


RESULT 11
FA7_RABIT
ID   FA7_RABIT      STANDARD;      PRT;   444 AA.
AC   P98139; P79224;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator).
GN   F7.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93190306; PubMed=8383365;
RA   Brothers A.B., Clarke B.J., Sheffield W.P., Blajchman M.A.;
RT   "Complete nucleotide sequence of the cDNA encoding rabbit coagulation
RT   factor VII.";
RL   Thromb. Res. Suppl. 69:231-238(1993).
RN   [2]
RP   REVISION TO 395.
RC   TISSUE=Liver;
RA   Ruiz S.R., Blajchman M.A., Clarke B.J.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Circulates in the blood in a zymogen form. Factor VII is
CC       converted to factor VIIa by factor Xa, factor XIIa, factor IXa, or
CC       thrombin by minor proteolysis. In the presence of tissue factor
CC       and calcium ions, factor VIIa then converts factor X to factor Xa
CC       by limited proteolysis. Factor VIIa will also convert factor IX to
CC       factor IXa in the presence of tissue factor and calcium (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC       form factor Xa.
CC   -!- SUBUNIT: Heterodimer of a light chain and a heavy chain linked by
CC       a disulfide bond (By similarity).
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U77477; AAB37326.1; -.
DR   HSSP; P08709; 1FAK.
DR   MEROPS; S01.215; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
```

```
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Blood coagulation; Zymogen; Glycoprotein;
KW   Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;
KW   EGF-like domain; Repeat; Signal; Hydroxylation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     39       POTENTIAL.
FT   CHAIN        40    191       FACTOR VII LIGHT CHAIN.
FT   CHAIN       192    444       FACTOR VII HEAVY CHAIN.
FT   DOMAIN       45     74       GLA-RICH.
FT   DOMAIN       85    121       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      126    167       EGF-LIKE 2.
FT   DOMAIN      192    444       SERINE PROTEASE.
FT   SITE        191    192       CLEAVAGE (BY FACTOR XA, FACTOR XIIA,
FT                                FACTOR IXA, OR THROMBIN) (BY SIMILARITY).
FT   ACT_SITE    232    232       BY SIMILARITY.
FT   ACT_SITE    281    281       BY SIMILARITY.
FT   ACT_SITE    383    383       BY SIMILARITY.
FT   BINDING     377    377       SUBSTRATE (BY SIMILARITY).
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    100       BY SIMILARITY.
FT   DISULFID     94    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    130    141       BY SIMILARITY.
FT   DISULFID    137    151       BY SIMILARITY.
FT   DISULFID    153    166       BY SIMILARITY.
FT   DISULFID    174    301       BY SIMILARITY.
FT   DISULFID    198    203       BY SIMILARITY.
FT   DISULFID    217    233       BY SIMILARITY.
FT   DISULFID    349    368       BY SIMILARITY.
FT   DISULFID    379    407       BY SIMILARITY.
FT   MOD_RES      45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     102    102       HYDROXYLATION (BY SIMILARITY).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   444 AA;  49011 MW;  0481ABC4FE5427F8 CRC64;

  Query Match             47.5%;  Score 94;  DB 1;  Length 444;
  Best Local Similarity   43.9%;  Pred. No. 6.5e-08;
  Matches   18;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              |||||  ||  || | |   :| |  |: :|:  : |  ||
Db         40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80


RESULT 12
FA7_HUMAN
ID   FA7_HUMAN      STANDARD;      PRT;   466 AA.
AC   P08709; Q14339;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator) (Eptacog alfa).
GN   F7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=86205965; PubMed=3486420;
RA   Hagen F.S., Gray C.L., O'Hara P.J., Grant F.J., Saari G.C.,
RA   Woodbury R.G., Hart C.E., Insley M.Y., Kisiel W., Kurachi K.,
RA   Davie E.W.;
RT   "Characterization of a cDNA coding for human factor VII.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:2412-2416(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87260948; PubMed=3037537;
RA   O'Hara P.J., Grant F.J., Haldeman B.A., Gray C.L., Insley M.Y.,
RA   Hagen F.S., Murray M.J.;
RT   "Nucleotide sequence of the gene coding for human factor VII, a
RT   vitamin K-dependent protein participating in blood coagulation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5158-5162(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS THR-352; GLN-413 AND LYS-445.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 61-466, AND POST-TRANSLATIONAL MODIFICATIONS.
RX   MEDLINE=89088153; PubMed=3264725;
RA   Thim L., Bjoern S., Christensen M., Nicolaisen E.M., Lund-Hansen T.,
RA   Pedersen A.H., Hedner U.;
RT   "Amino acid sequence and posttranslational modifications of human
RT   factor VIIa from plasma and transfected baby hamster kidney cells.";
RL   Biochemistry 27:7785-7793(1988).
RN   [5]
RP   CARBOHYDRATE-LINKAGE SITES SER-112 AND SER-120.
RX   MEDLINE=91250411; PubMed=1904059;
RA   Bjoern S., Foster D.C., Thim L., Wiberg F.C., Christensen M.,
RA   Komiyama Y., Pedersen A.H., Kisiel W.;
RT   "Human plasma and recombinant factor VII. Characterization of O-
RT   glycosylations at serine residues 52 and 60 and effects of site-
RT   directed mutagenesis of serine 52 to alanine.";
RL   J. Biol. Chem. 266:11051-11057(1991).
RN   [6]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=90062160; PubMed=2511201;
RA   Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T., Takao T.,
RA   Shimonishi Y., Iwanaga S.;
RT   "Identification of a disaccharide (Xyl-Glc) and a trisaccharide
RT   (Xyl2-Glc) O-glycosidically linked to a serine residue in the first
RT   epidermal growth factor-like domain of human factors VII and IX and
RT   protein Z and bovine protein Z.";
RL   J. Biol. Chem. 264:20320-20325(1989).
RN   [7]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=91344709; PubMed=2129367;
RA   Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;
RT   "A new trisaccharide sugar chain linked to a serine residue in the
```

```
RT   first EGF-like domain of clotting factors VII and IX and protein Z.";
RL   Adv. Exp. Med. Biol. 281:121-131(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=96175641; PubMed=8598903;
RA   Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,
RA   Konigsberg W.H., Nemreson Y., Kirchhofer D.;
RT   "The crystal structure of the complex of blood coagulation factor
RT   VIIa with soluble tissue factor.";
RL   Nature 380:41-46(1996).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=99126538; PubMed=9925787;
RA   Zhang E., St Charles R., Tulinsky A.;
RT   "Structure of extracellular tissue factor complexed with factor VIIa
RT   inhibited with a BPTI mutant.";
RL   J. Mol. Biol. 285:2089-2104(1999).
RN   [10]
RP   STRUCTURE BY NMR OF 105-145.
RX   MEDLINE=98367502; PubMed=9692950;
RA   Muranyi A., Finn B.E., Gippert G.P., Forsen S., Stenflo J.,
RA   Drakenberg T.;
RT   "Solution structure of the N-terminal EGF-like domain from human
RT   factor VII.";
RL   Biochemistry 37:10605-10615(1998).
RN   [11]
RP   VARIANT GLN-364.
RX   MEDLINE=91300046; PubMed=2070047;
RA   O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,
RA   Meade T.W., Tuddenham E.G.D.;
RT   "Purification and characterization of factor VII 304-Gln: a variant
RT   molecule with reduced activity isolated from a clinically unaffected
RT   male.";
RL   Blood 78:132-140(1991).
RN   [12]
RP   VARIANTS GLN-364 AND PHE-370.
RX   MEDLINE=92340074; PubMed=1634227;
RA   Marchetti G., Patracchini P., Gemmati D., Derosa V., Pinotti M.,
RA   Rodorigo G., Casonato A., Girolami A., Bernardi F.;
RT   "Detection of two missense mutations and characterization of a repeat
RT   polymorphism in the factor VII gene (F7).";
RL   Hum. Genet. 89:497-502(1992).
RN   [13]
RP   VARIANT TYR-238.
RX   MEDLINE=93372811; PubMed=8364544;
RA   Marchetti G., Ferrati M., Patracchini P., Redaelli R., Bernardi F.;
RT   "A missense mutation (178Cys-->Tyr) and two neutral dimorphisms
RT   (115His and 333Ser) in the human coagulation factor VII gene.";
RL   Hum. Mol. Genet. 2:1055-1056(1993).
RN   [14]
RP   VARIANTS.
RX   MEDLINE=94061028; PubMed=8242057;
RA   Takamiya O., Kemball-Cook G., Martin D.M.A., Cooper D.N.,
RA   von Felten A., Meili E., Hahn I., Prangnell D.R., Lumley H.,
RA   Tuddenham E.G.D., McVey J.H.;
RT   "Detection of missense mutations by single-strand conformational
RT   polymorphism (SSCP) analysis in five dysfunctional variants of
RT   coagulation factor VII.";
RL   Hum. Mol. Genet. 2:1355-1359(1993).
RN   [15]
RP   VARIANTS CHARLOTTE GLN-139 AND GLN-212.
RX   MEDLINE=94264305; PubMed=8204879;
RA   Chaing S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,
RA   Roberts H.R., Blajchman M., Monroe D.M., High K.A.;
RT   "Severe factor VII deficiency caused by mutations abolishing the
RT   cleavage site for activation and altering binding to tissue factor.";
RL   Blood 83:3524-3535(1994).
RN   [16]
RP   VARIANT VAL-354.
RX   MEDLINE=95072589; PubMed=7981691;
RA   Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,
RA   Rodeghiero F., Marchetti G.;
RT   "Topologically equivalent mutations causing dysfunctional coagulation
RT   factors VII (294Ala-->Val) and X (334Ser-->Pro).";
RL   Hum. Mol. Genet. 3:1175-1177(1994).
RN   [17]
RP   VARIANT MIE HIS-307.
RX   MEDLINE=95064662; PubMed=7974346;
RA   Ohiwa M., Hayashi T., Wada H., Minamikawa K., Shirakawa S.,
RA   Suzuki K.;
RT   "Factor VII Mie: homozygous asymptomatic type I deficiency caused by
RT   an amino acid substitution of His (CAC) for Arg(247) (CGC) in the
RT   catalytic domain.";
RL   Thromb. Haemost. 71:773-777(1994).
RN   [18]
RP   VARIANT MET-419.
RX   MEDLINE=96247510; PubMed=8652821;
RA   Arbini A.A., Mannucci P.M., Bauer K.A.;
RT   "A Thr359Met mutation in factor VII of a patient with a hereditary
RT   deficiency causes defective secretion of the molecule.";
RL   Blood 87:5085-5094(1996).
RN   [19]
RP   VARIANTS TRP-283; LYS-325; VAL-358; GLN-364; GLU-402 AND GLN-413.
RX   MEDLINE=97001216; PubMed=8844208;
RA   Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,
RA   Lunghi B., Rodeghiero F., Marchetti G.;
RT   "Mutation pattern in clinically asymptomatic coagulation factor VII
RT   deficiency.";
RL   Hum. Mutat. 8:108-115(1996).
RN   [20]
RP   VARIANT VAL-304.
RX   MEDLINE=97037613; PubMed=8883260;
RA   Tamary H., Fromovich Y., Shalmon L., Reich Z., Dym O., Lanir N.,
RA   Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,
RA   Zaizov R., Seligsohn U.;
RT   "Ala244Val is a common, probably ancient mutation causing factor VII
RT   deficiency in Moroccan and Iranian Jews.";
RL   Thromb. Haemost. 76:283-291(1996).
RN   [21]
RP   VARIANT MORIOKA PRO-13.
RX   MEDLINE=98235713; PubMed=9576180;
RA   Ozawa T., Takikawa Y., Niiya K., Ejiri N., Suzuki K., Sato S.,
RA   Sakuragawa N.;
RT   "Factor VII Morioka (FVII L-26P): a homozygous missense mutation in
RT   the signal sequence identified in a patient with factor VII
RT   deficiency.";
RL   Br. J. Haematol. 101:47-49(1998).
RN   [22]
RP   VARIANTS MALTA THR-194 AND VAL-304.
RX   MEDLINE=98112461; PubMed=9452082;
RA   Alshinawi C., Scerri C., Galdies R., Aquilina A., Felice A.E.;
RT   "Two new missense mutations (P134T and A244V) in the coagulation
RT   factor VII gene.";
RL   Hum. Mutat. Suppl. 1:S189-S191(1998).
RN   [23]
RP   VARIANTS ASP-295 AND GLN-413.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).

  Query Match             46.5%;  Score 92;  DB 1;  Length 466;
  Best Local Similarity   46.3%;  Pred. No. 1.5e-07;
  Matches   19;  Conservative    4;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||:||  ||  || | |    | |  |: ||:| : |  ||
Db         61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 13
TMG1_HUMAN
```

```
ID   TMG1_HUMAN     STANDARD;      PRT;   218 AA.
AC   O14668;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 1 precursor (Proline-
DE   rich Gla protein 1) (Proline-rich gamma-carboxyglutamic acid protein
DE   1).
GN   PRRG1 OR TMG1 OR PRGP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97404347; PubMed=9256434;
RA   Kulman J.D., Harris J.E., Haldeman B.A., Davie E.W.;
RT   "Primary structure and tissue distribution of two novel proline-rich
RT   gamma-carboxyglutamic acid proteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9058-9062(1997).
CC   -!- TISSUE SPECIFICITY: Highly expressed in the spinal cord.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamate by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF009242; AAB67070.1; -.
DR   HSSP; P00740; 1CFH.
DR   Genew; HGNC:9469; PRRG1.
DR   MIM; 604428; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     20       POTENTIAL.
FT   CHAIN          21    218       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 1.
FT   DOMAIN         21     83       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       84    106       POTENTIAL.
FT   DOMAIN        107    218       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         24     61       GLA-RICH.
FT   DOMAIN        131    135       POLY-PRO.
SQ   SEQUENCE   218 AA;  24947 MW;  26538A61AB0AEB98 CRC64;

  Query Match             40.9%;  Score 81;  DB 1;  Length 218;
  Best Local Similarity   34.1%;  Pred. No. 4.4e-06;
  Matches   15;  Conservative    7;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              || |   :|  :: | |     | |  |:  ||: : |  ||| :
Db         21 ANGFFEEIRQGNIERECKEEFCTFEEAREAFENNEKTKEFWSTY 64


RESULT 14
THRB_RAT
ID   THRB_RAT       STANDARD;      PRT;   617 AA.
AC   P18292;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=90332426; PubMed=2377469;
RA   Dihanich M., Monard D.;
RT   "cDNA sequence of rat prothrombin.";
RL   Nucleic Acids Res. 18:4251-4251(1990).
RN   [2]
RP   SEQUENCE OF 383-617 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
CC       fragment (fragment 1) of the prothrombin, prior to its activation
CC       by factor Xa.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52835; CAA37017.1; -.
DR   EMBL; M81397; AAA42240.1; -.
DR   PIR; S10511; S10511.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR003966; Peptidase_S1A_pr.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
```

```
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   PROPEP       25     43
FT   CHAIN        44    617       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    323       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       324    359       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       360    617       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      360    617       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        323    324       CLEAVAGE (BY FACTOR XA).
FT   SITE        359    360       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    402    402       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    458    458       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    564    564       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD    120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    552    552       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    292       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    287       BY SIMILARITY.
FT   DISULFID    332    478       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    387    403       BY SIMILARITY.
FT   DISULFID    532    546       BY SIMILARITY.
FT   DISULFID    560    590       BY SIMILARITY.
SQ   SEQUENCE   617 AA;  70411 MW;  AD27D1B71445DB1D CRC64;

  Query Match             39.6%;  Score 78.5;  DB 1;  Length 617;
  Best Local Similarity   40.0%;  Pred. No. 3.7e-05;
  Matches   18;  Conservative    5;  Mismatches   21;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||  :| | |:   | :  |    |   ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 15
THRB_MOUSE
ID   THRB_MOUSE     STANDARD;      PRT;   618 AA.
AC   P19221;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2 OR CF2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=91025551; PubMed=2222810;
RA   Friezner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA   Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse prothrombin and
RT   localization of the gene on mouse chromosome 2.";
RL   DNA Cell Biol. 9:487-498(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 384-618 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
CC       converts fibrinogen to fibrin and activates factors V, VII, VIII,
CC       XIII, and, in complex with thrombomodulin, protein C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: Prothrombin is activated on the surface of a
CC       phospholipid membrane that binds the amino end of prothrombin and
CC       factors Va and Xa in Ca-dependent interactions; factor Xa removes
CC       the activation peptide and cleaves the remaining part into light
CC       and heavy chains. The activation process starts slowly because
CC       factor V itself has to be activated by the initial, small amounts
CC       of thrombin.
CC   -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
CC       fragment (fragment 1) of the prothrombin, prior to its activation
CC       by factor Xa.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X52308; CAA36548.1; -.
DR   EMBL; BC013662; AAH13662.1; -.
DR   EMBL; M81394; AAA40435.1; -.
DR   PIR; A35827; A35827.
DR   HSSP; P00734; 1B7X.
DR   MEROPS; S01.217; -.
DR   MGD; MGI:88380; F2.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR003966; Peptidase_S1A_pr.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   PROPEP      25     43
FT   CHAIN        44    618       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    324       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       325    360       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       361    618       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      361    618       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        324    325       CLEAVAGE (BY FACTOR XA).
FT   SITE        360    361       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    403    403       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    459    459       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    565    565       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    293       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    288       BY SIMILARITY.
```

```
FT   DISULFID    333    479
FT   DISULFID    388    404
FT   DISULFID    533    547
FT   DISULFID    561    591       E
FT   CARBOHYD    122    122       N
FT   CARBOHYD    144    144       N-
FT   CARBOHYD    413    413       N-
FT   CARBOHYD    553    553       N-L
SQ   SEQUENCE   618 AA;  70268 MW;  B

  Query Match           39.6%;  Score
  Best Local Similarity  40.0%;  Pred.
  Matches   18;  Conservative    5;  Mis

Qy          1 ANS-FLXXLRHSSLXRXCIXXICDFXX
              |||  ||  ||  :|  |  |:   | :  |
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEAL

Search completed: March  1, 2004, 10:03:23
Job time : 11.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:55:12 ; Search time 37.5 Seconds
(without alignments)
370.208 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	144	72.7	456	6	Q9TTR0	Q9ttr0 canis famil
2	133	67.2	460	11	Q91WN8	Q91wn8 mus musculu
3	130	65.7	55	4	Q8J002	Q8j002 homo sapien
4	130	65.7	55	4	Q8IXB5	Q8ixb5 homo sapien
5	127	64.1	460	11	Q99PC6	Q99pc6 mus musculu
6	107	54.0	482	11	Q63207	Q63207 rattus norv
7	102	51.5	455	13	Q7SY86	Q7sy86 xenopus lae
8	100	50.5	433	13	Q804X5	Q804x5 gallus gall
9	98	49.5	524	13	Q7SXH8	Q7sxh8 brachydanio
10	96.5	48.7	443	13	Q8JHC9	Q8jhc9 brachydanio
11	96	48.5	231	4	Q8N2N6	Q8n2n6 homo sapien
12	93	47.0	340	11	Q80Y26	Q80y26 mus musculu
13	93	47.0	434	13	Q7T3B6	Q7t3b6 brachydanio
14	93	47.0	481	11	O54740	O54740 mus musculu
15	93	47.0	481	11	Q99L32	Q99132 mus musculu
16	93	47.0	481	11	O88947	O88947 mus musculu
17	92	46.5	679	4	Q96PQ8	Q96pq8 homo sapien
18	89	44.9	474	13	Q8JHC8	Q8jhc8 brachydanio
19	88.5	44.7	442	13	Q804X1	Q804x1 fugu rubrip
20	87	43.9	469	6	Q9GMD9	Q9gmd9 ornithorhyn
21	82	41.4	229	13	Q8JJ40	Q8jj40 xenopus lae
22	81	40.9	268	4	Q8NEK6	Q8nek6 homo sapien
23	80	40.4	425	13	Q804X7	Q804x7 gallus gall
24	80	40.4	612	13	Q804W7	Q804w7 fugu rubrip
25	79	39.9	446	11	Q8K3U6	Q8k3u6 rattus norv
26	76	38.4	100	4	Q15253	Q15253 homo sapien
27	76	38.4	622	4	Q7Z7P3	Q7z7p3 homo sapien
28	74	37.4	471	13	Q804X6	Q804x6 gallus gall
29	74	37.4	475	13	Q804W9	Q804w9 fugu rubrip
30	74	37.4	497	4	Q7Z715	Q7z715 homo sapien
31	74	37.4	650	4	Q16519	Q16519 homo sapien
32	74	37.4	650	4	Q9NSD0	Q9nsd0 homo sapien
33	73	36.9	446	11	Q61109	Q61109 mus musculu
34	71.5	36.1	542	5	Q8T6I3	Q8t6i3 halocynthia
35	71	35.9	138	6	Q28994	Q28994 sus scrofa
36	71	35.9	441	13	Q804X2	Q804x2 fugu rubrip
37	70	35.4	607	13	Q91001	Q91001 gallus gall
38	69	34.8	461	6	Q95ND7	Q95nd7 pan troglod
39	69	34.8	461	6	Q95ND6	Q95nd6 pan troglod
40	67	33.8	648	6	Q29094	Q29094 sus scrofa
41	66.5	33.6	433	13	Q8JHD0	Q8jhd0 brachydanio
42	66.5	33.6	433	13	Q90YK1	Q90yk1 brachydanio
43	66	33.3	49	6	Q95ME8	Q95me8 bos taurus
44	66	33.3	52	4	Q8IXD5	Q8ixd5 homo sapien
45	64	32.3	430	13	Q804X0	Q804x0 fugu rubrip

ALIGNMENTS

RESULT 1
Q9TTR0
ID Q9TTR0 PRELIMINARY; PRT; 456 AA.
AC Q9TTR0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein C precursor.
GN PROC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Leeb T., Kopp T., Deppe A., Breen M., Matis U., Brunnberg L.,
RA Brenig B.;
RT "Molecular characterization and chromosomal assignment of the canine
RT protein C gene.";
RL Mamm. Genome 10:135-139(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99371952; PubMed=10443005;
RA Leeb T., Pfeiffer I., Kopp T., Deppe A., Brenig B.;
RT "Analysis of canine protein C gene polymorphisms.";
RL Anim. Genet. 30:237-238(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AJ001979; CAA05126.1; -.
DR HSSP; P04070; 1AUT.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.

```
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease; Signal.
FT   SIGNAL        1     42       POTENTIAL.
FT   CHAIN        43    192       PROTEIN C LIGHT CHAIN.
FT   CHAIN       193    194       PROTEIN C CONNECTING DIPEPTIDE.
FT   CHAIN       195    456       PROTEIN C HEAVY CHAIN.
SQ   SEQUENCE   456 AA;  50813 MW;  7AD3A8C1C34E59FF CRC64;

  Query Match             72.7%;  Score 144;  DB 6;  Length 456;
  Best Local Similarity   61.4%;  Pred. No. 9.3e-16;
  Matches   27;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  |||| || ||::||||||:|||:
Db         43 ANSFLEEIRAGSLERECMEEICDFEEAKEIFQNVDDTLAYWSKY 86


RESULT 2
Q91WN8
ID   Q91WN8      PRELIMINARY;      PRT;   460 AA.
AC   Q91WN8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; BC013896; AAH13896.1; -.
DR   HSSP; P00761; 1AN1.
DR   MGD; MGI:97771; Proc.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   460 AA;  51818 MW;  0117F26E68FCC274 CRC64;

  Query Match             67.2%;  Score 133;  DB 11;  Length 460;
  Best Local Similarity   56.8%;  Pred. No. 6.9e-14;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  |||| |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
Q8J002
ID   Q8J002      PRELIMINARY;      PRT;   55 AA.
AC   Q8J002;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein C (Fragment).
GN   PROC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kinoshita S., Iida H., Inoue S., Watanabe K., Kurihara M., Wada Y.,
RA   Ono M., Dongchon K., Hamasaki N.;
RT   "Gene Analysis of Anticoagulation Factors in Japanese Thrombotic
RT   Patients. Genetic Background of Thrombophilia in Japan.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB083700; BAC21172.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
FT   NON_TER       1      1
FT   NON_TER      55     55
SQ   SEQUENCE   55 AA;  6527 MW;  4F89496534A78836 CRC64;

  Query Match             65.7%;  Score 130;  DB 4;  Length 55;
  Best Local Similarity   70.3%;  Pred. No. 2.2e-14;
  Matches   26;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDT 37
              |||||  |||||| | ||  |||| || ||::||||
Db         19 ANSFLKELRHSSLERECIEEICDFEEAKEIFQNVDDT 55


RESULT 4
Q8IXB5
ID   Q8IXB5      PRELIMINARY;      PRT;   55 AA.
AC   Q8IXB5;
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein C (Fragment).
GN   PROC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hamasaki S., Kang D., Kinoshita S., Iida K., Inoue S., Watanabe K.,
RA   Kurihara M., Wada Y., Ono M.;
RT   "Gene analysis of anticoagulation factors in Japanese thrombotic
RT   patients.Genetic background of thrombophilia in Japan.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB086851; BAC53631.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
FT   NON_TER       1      1
FT   NON_TER      55     55
SQ   SEQUENCE   55 AA;  6475 MW;  3803696534BC9289 CRC64;

  Query Match             65.7%;  Score 130;  DB 4;  Length 55;
  Best Local Similarity   70.3%;  Pred. No. 2.2e-14;
  Matches   26;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDT 37
              |||||  ||||||  | ||  ||||  || ||::|||
Db         19 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDT 55


RESULT 5
Q99PC6
ID   Q99PC6      PRELIMINARY;      PRT;   460 AA.
AC   Q99PC6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Anticoagulant protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL;
RA   Korf I.;
RT   "Complete sequence of UC72A01.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF318182; AAK07918.1; -.
DR   HSSP; P04070; 1AUT.
DR   MGD; MGI:97771; Proc.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   460 AA;  51784 MW;  0293BC25E9D3ED16 CRC64;

  Query Match             64.1%;  Score 127;  DB 11;  Length 460;
  Best Local Similarity   54.5%;  Pred. No. 7.2e-13;
  Matches   24;  Conservative    7;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :|  || | |:  |||   |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDLEEAQEIFQNVEDTLAFWIKY 85


RESULT 6
Q63207
ID   Q63207      PRELIMINARY;      PRT;   482 AA.
AC   Q63207;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Factor X.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=96093366; PubMed=8578539;
RA   Stanton C., Ross R.P., Hutson S., Wallin R.;
RT   "Evidence for competition between vitamin K-dependent clotting factors
RT   for intracellular processing by the vitamin K-dependent gamma-
RT   carboxylase.";
RL   Thromb. Res. 80:63-73(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; X79807; CAA56202.1; -.
DR   PIR; S49075; EXRT.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
```

```
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   482 AA;  54265 MW;  0284678E3954A698 CRC64;

  Query Match             54.0%;  Score 107;  DB 11;  Length 482;
  Best Local Similarity   40.9%;  Pred. No. 1.9e-09;
  Matches   18;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||    ::  :| | |:  || |  |: :||| : |  ||:|:
Db         41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 7
Q7SY86
ID   Q7SY86      PRELIMINARY;      PRT;   455 AA.
AC   Q7SY86;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole;
RA   Klein S., Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054968; AAH54968.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   455 AA;  51811 MW;  07C027ED2B495330 CRC64;

  Query Match             51.5%;  Score 102;  DB 13;  Length 455;
  Best Local Similarity   47.7%;  Pred. No. 1.3e-08;
  Matches   21;  Conservative    6;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              | :|:  |:  || | ||   |||  |  |||  :||| ||:|:
Db         49 AFNFMEELKPGSLERECIEEKCDFEEAFEIFETKEDTLNFWAKY 92


RESULT 8
Q804X5
ID   Q804X5      PRELIMINARY;      PRT;   433 AA.
AC   Q804X5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Anticoagulant protein C precursor (EC 3.4.21.69).
GN   PROC.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RT   "Comparative sequence analysis and molecular evolution of blood
RT   coagulation genes from Gallus gallus and Fugu rubripes.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF465270; AAO33365.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003808; F:protein C (activated) activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
```

```
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase.
SQ   SEQUENCE   433 AA;  48689 MW;  E09DDEE56D7DA2A3 CRC64;

  Query Match             50.5%;  Score 100;  DB 13;  Length 433;
  Best Local Similarity   47.7%;  Pred. No. 2.6e-08;
  Matches   21;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |:  |: | |     |:|  |  |||  : || ||||:
Db         40 ANSFLEELKPGSVERECNEERCNFEEASEIFETKEATLEFWSKY 83


RESULT 9
Q7SXH8
ID   Q7SXH8      PRELIMINARY;      PRT;   524 AA.
AC   Q7SXH8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC055596; AAH55596.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   524 AA;  59560 MW;  1B4AE08119080325 CRC64;

  Query Match             49.5%;  Score 98;  DB 13;  Length 524;
  Best Local Similarity   40.9%;  Pred. No. 7e-08;
  Matches   18;  Conservative    8;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||:    |:  :| | |:  |||   |: :|| || |  ||:|:
Db         42 ANTVFEELKPGNLERECVEEICDHEEAREVFERVDKTEIFWAKY 85


RESULT 10
Q8JHC9
ID   Q8JHC9      PRELIMINARY;      PRT;   443 AA.
AC   Q8JHC9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor VIIi.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hanumanthaiah R., Day K., Jagadeeswaran P.;
RT   "Comprehensive analysis of blood coagulation pathways in Teleostei:
RT   Evolution of coagulation factor genes and identification of zebrafish
RT   factor VIIi.";
RL   Blood Cells Mol. Dis. 0:0-0(2002).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF519546; AAM88342.1; -.
DR   EMBL; AF515269; AAN71000.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   443 AA;  48823 MW;  2D2504718AE94FF4 CRC64;

  Query Match             48.7%;  Score 96.5;  DB 13;  Length 443;
  Best Local Similarity   45.2%;  Pred. No. 1.1e-07;
  Matches   19;  Conservative    7;  Mismatches   15;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||| ||  ::  :| | |:  |||:  |: :||| | |  ||
Db         38 ANSGFLEEMKAGNLERECVEEICDYEEAREVFEDDDRTKQFW 79


RESULT 11
Q8N2N6
ID   Q8N2N6      PRELIMINARY;      PRT;   231 AA.
AC   Q8N2N6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein FLJ90093.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA   Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA   Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA   Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK074574; BAC11069.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   231 AA;  25844 MW;  8A373B0D5C1D0D81 CRC64;

  Query Match             48.5%;  Score 96;  DB 4;  Length 231;
  Best Local Similarity   39.0%;  Pred. No. 6.3e-08;
  Matches   16;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  ||  :: | |:  || :   | :||: : |: ||
Db         20 ANEFLEELRQGTIERECMEEICSYEEVKEVFENKEKTMEFW 60


RESULT 12
Q80Y26
ID   Q80Y26      PRELIMINARY;      PRT;   340 AA.
AC   Q80Y26;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   F10 protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RA   Strausberg R.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC050219; AAH50219.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
SQ   SEQUENCE   340 AA;  38359 MW;  EE252D6157720811 CRC64;

  Query Match             47.0%;  Score 93;  DB 11;  Length 340;
  Best Local Similarity   36.4%;  Pred. No. 3.1e-07;
  Matches   16;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||    :  :| | |:  || :   : |||| : |  :|:|:
Db         53 ANSFPEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 96


RESULT 13
Q7T3B6
ID   Q7T3B6      PRELIMINARY;      PRT;   434 AA.
AC   Q7T3B6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053182; AAH53182.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   434 AA;  48516 MW;  B47BD7947CF9D9C9 CRC64;

  Query Match             47.0%;  Score 93;  DB 13;  Length 434;
  Best Local Similarity   40.9%;  Pred. No. 4e-07;
  Matches   18;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||:|:  |: :|| | |   :|||  |: ||   : || ||: :
Db         40 ANTFMEELKPASLERECREELCDFEEAREIFITREATLEFWTAY 83


RESULT 14
O54740
ID   O54740      PRELIMINARY;      PRT;   481 AA.
AC   O54740;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6).
GN   F10 OR FA10.
OS   Mus musculus (Mouse).
OG   Plasmid pBluescript.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=98454993; PubMed=9783672;
RA   Heidtmann H.H., Kontermann R.E.;
RT   "Cloning and recombinant expression of mouse coagulation factor X.";
RL   Thromb. Res. 92:33-41(1998).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AJ222677; CAA10933.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0003804; F:blood coagulation factor X activity; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease; Signal;
KW   Plasmid.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41    481       COAGULATION FACTOR X.
SQ   SEQUENCE   481 AA;  53986 MW;  CF702DE5EF9D97AE CRC64;

  Query Match             47.0%;  Score 93;  DB 11;  Length 481;
  Best Local Similarity   36.4%;  Pred. No. 4.5e-07;
  Matches   16;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||     :  :| | |:  || :   : |||| : |  :|:|:
Db         41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 15
Q99L32
ID   Q99L32      PRELIMINARY;      PRT;   481 AA.
AC   Q99L32;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coagulation factor X.
GN   F10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; BC003877; AAH03877.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
```

```
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   EGF-like domain; Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   481 AA;  54004 MW;  BD88E96C8A0B7E7F CRC64;

  Query Match             47.0%;  Score 93;  DB 11;  Length 481;
  Best Local Similarity   36.4%;  Pred. No. 4.5e-07;
  Matches   16;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||     :  :| | |:  || :   : |||| : |  :|:|:
Db         41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


Search completed: March  1, 2004, 10:02:51
Job time : 40.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 10:03:28 ; Search time 16.5 Seconds
(without alignments)
137.669 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	180	90.9	44	3	US-08-955-636-19	Sequence 19, Appl
2	176	88.9	419	4	US-10-182-263-3	Sequence 3, Appli
3	176	88.9	419	4	US-10-182-263-4	Sequence 4, Appli
4	176	88.9	419	4	US-10-182-263-5	Sequence 5, Appli
5	172	86.9	44	3	US-08-955-636-1	Sequence 1, Appli
6	172	86.9	44	3	US-08-955-636-20	Sequence 20, Appl
7	172	86.9	44	3	US-08-955-636-21	Sequence 21, Appl
8	172	86.9	44	3	US-08-955-636-25	Sequence 25, Appl
9	172	86.9	45	2	US-08-965-832-2	Sequence 2, Appli
10	172	86.9	419	1	US-08-295-411-1	Sequence 1, Appli
11	172	86.9	419	2	US-08-955-471-1	Sequence 1, Appli
12	172	86.9	419	4	US-09-667-570A-3	Sequence 3, Appli
13	172	86.9	419	4	US-10-182-263-1	Sequence 1, Appli
14	172	86.9	419	5	PCT-US92-10242-1	Sequence 1, Appli
15	172	86.9	460	2	US-08-756-506-2	Sequence 2, Appli
16	172	86.9	460	2	US-08-756-506-4	Sequence 4, Appli
17	172	86.9	460	6	5270178-13	Patent No. 5270178
18	172	86.9	460	6	5270178-14	Patent No. 5270178
19	172	86.9	460	6	5270178-15	Patent No. 5270178
20	172	86.9	460	6	5270178-16	Patent No. 5270178
21	172	86.9	461	4	US-10-182-263-2	Sequence 2, Appli
22	172	86.9	461	6	5225537-2	Patent No. 5225537
23	172	86.9	461	6	5270178-17	Patent No. 5270178
24	172	86.9	461	6	5270178-18	Patent No. 5270178
25	172	86.9	461	6	5460953-3	Patent No. 5460953
26	171	86.4	44	3	US-08-955-636-22	Sequence 22, Appl
27	168	84.8	44	3	US-08-955-636-24	Sequence 24, Appl
28	168	84.8	44	3	US-08-955-636-35	Sequence 35, Appl
29	168	84.8	419	4	US-10-182-263-6	Sequence 6, Appli
30	159	80.3	42	2	US-08-745-254A-2	Sequence 2, Appli
31	159	80.3	461	6	5270178-2	Patent No. 5270178
32	155	78.3	41	1	US-08-229-280-5	Sequence 5, Appli
33	150	75.8	42	4	US-09-383-667-8	Sequence 8, Appli
34	141	71.2	410	3	US-09-065-872-1	Sequence 1, Appli
35	141	71.2	410	4	US-09-667-570A-1	Sequence 1, Appli
36	133	67.2	409	3	US-09-065-872-2	Sequence 2, Appli
37	133	67.2	409	4	US-09-667-570A-2	Sequence 2, Appli
38	119	60.1	44	3	US-08-955-636-23	Sequence 23, Appl
39	109	55.1	44	3	US-08-955-636-2	Sequence 2, Appli
40	106	53.5	139	1	US-08-330-978-2	Sequence 2, Appli
41	106	53.5	139	1	US-08-474-042-2	Sequence 2, Appli
42	106	53.5	139	1	US-08-484-558-2	Sequence 2, Appli
43	106	53.5	139	1	US-08-774-592-2	Sequence 2, Appli
44	106	53.5	437	1	US-08-487-037-2	Sequence 2, Appli
45	106	53.5	437	1	US-08-487-037-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-955-636-19
; Sequence 19, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-19

Query Match 90.9%; Score 180; DB 3; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||||||||||||||||||||||||||||||||||||||||||
Db 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44

RESULT 2
US-10-182-263-3
; Sequence 3, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-3

Query Match 88.9%; Score 176; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 9.3e-22;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||| ||| || | || |||| || ||||||||||||||||
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 3
US-10-182-263-4
; Sequence 4, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-4

Query Match 88.9%; Score 176; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 9.3e-22;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||| ||| || | || |||| || ||||||||||||||||
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 4
US-10-182-263-5
; Sequence 5, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-5

Query Match 88.9%; Score 176; DB 4; Length 419;
Best Local Similarity 77.3%; Pred. No. 9.3e-22;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||| ||| || | || |||| || ||||||||||||||||
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 5
US-08-955-636-1
; Sequence 1, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-1

Query Match 86.9%; Score 172; DB 3; Length 44;
Best Local Similarity 95.5%; Pred. No. 3.3e-22;
Matches 42; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
|||||||||||||||||||||||||||||::||||||||||
Db 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44

RESULT 6
US-08-955-636-20
; Sequence 20, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-20

Query Match 86.9%; Score 172; DB 3; Length 44;
Best Local Similarity 97.7%; Pred. No. 3.3e-22;
Matches 43; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||||||| |||||||||||||||||||||||||||||||||||
Db 1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44

RESULT 7

```
)8-955-636-21
:quence 21, Application US/08955636A
itent No. 6017882
ENERAL INFORMATION:
APPLICANT: Nelsestuen, Gary
ITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
ITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 09531/002001
CURRENT APPLICATION NUMBER: US/08/955,636A
CURRENT FILING DATE: 1997-10-23
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSEQ for Windows Version 3.0
EQ ID NO 21
 LENGTH: 44
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: MOD_RES
 LOCATION: (0)...(0)
 OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
-08-955-636-21

 Query Match             86.9%;  Score 172;  DB 3;  Length 44;
 Best Local Similarity   97.7%;  Pred. No. 3.3e-22;
 Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

y          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
             ||||||||| |||||||||||||||||||||||||||||||||||
b          1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


ESULT 8
S-08-955-636-25
 Sequence 25, Application US/08955636A
 Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-25

  Query Match             86.9%;  Score 172;  DB 3;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 3.3e-22;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
             |||||||||||||||||||||||||| ||||||||||||||||||
Db         1 ANSFLXXLRHSSLXRXCIXXICDFXXAFXIFEDVDDTLAFWSKH 44


RESULT 9
US-08-965-832-2
; Sequence 2, Application US/08965832
; Patent No. 5847085
;  GENERAL INFORMATION:
;    APPLICANT:  CHARLES T. ESMON AND MIKHAIL D. SMIRNOV
;    TITLE OF INVENTION:  Modified Protein C
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center, 1201 West
;      STREET:  Peachtree Street
;      CITY:  Atlanta
;      STATE:  GA
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/965,832
;      FILING DATE:  7-NOV-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/745,254
;      FILING DATE:  8-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/053,768
;      FILING DATE:  25-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF 165/167
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-873-8794
;      TELEFAX:  (404)-873-8795
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:
;      LOCATION:  6, 7, 14, 16, 19, 20, 25, 26, 29
;      OTHER INFORMATION:  /note= "where Xaa means gamma
;      OTHER INFORMATION:  carboxylglutamic acid"
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      OTHER INFORMATION:  /note= "partial sequence of human protein C"
US-08-965-832-2

  Query Match             86.9%;  Score 172;  DB 2;  Length 45;
  Best Local Similarity   95.5%;  Pred. No. 3.3e-22;
  Matches   42;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
             ||||||||||||||||||||||||||||||::|||||||||||
Db         1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 10
US-08-295-411-1
; Sequence 1, Application US/08295411
; Patent No. 5679639
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5679639th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
```

```
      ZIP:  92037
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/295,411
      FILING DATE:  22-AUG-1994
      CLASSIFICATION:  530
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/793,989
      FILING DATE:  18-NOV-1991
      CLASSIFICATION:  530
    ATTORNEY/AGENT INFORMATION:
      NAME:  Fitting, Thomas
      REGISTRATION NUMBER:  34,163
      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  619-554-2937
      TELEFAX:  619-554-6312
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  419 amino acids
      TYPE:  amino acid
      TOPOLOGY:  linear
    MOLECULE TYPE:  protein
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    FEATURE:
      NAME/KEY:  Region
      LOCATION:  1..157
      OTHER INFORMATION:  /note= "Protein C Light Chain"
    FEATURE:
      NAME/KEY:  Region
      LOCATION:  158..169
      OTHER INFORMATION:  /note= "Protein C Activation
      OTHER INFORMATION:  Peptide"
    FEATURE:
      NAME/KEY:  Region
      LOCATION:  170..419
      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-295-411-1

  Query Match             86.9%;  Score 172;  DB 1;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 4.5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 11
US-08-955-471-1
; Sequence 1, Application US/08955471
; Patent No. 5968751
; GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5968751th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,471
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/295,411
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  419 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..157
;      OTHER INFORMATION:  /note= "Protein C Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  158..169
;      OTHER INFORMATION:  /note= "Protein C Activation
;      OTHER INFORMATION:  Peptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  170..419
;      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-955-471-1

  Query Match             86.9%;  Score 172;  DB 2;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 4.5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 12
US-09-667-570A-3
; Sequence 3, Application US/09667570A
; Patent No. 6436397
; GENERAL INFORMATION:
;  APPLICANT: Baker, Jeffrey C
;  APPLICANT:  Carlson, Andrew D
;  APPLICANT:  Huang, Lihua
;  APPLICANT:  Sheliga, Theodore  A
;  TITLE OF INVENTION: Improved Methods for Processing Activated Protein C
;  FILE REFERENCE: X-11796A
;  CURRENT APPLICATION NUMBER: US/09/667,570A
;  CURRENT FILING DATE:  2000-09-21
;  PRIOR APPLICATION NUMBER: 60/045,255
;  PRIOR FILING DATE: 1997-04-28
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;    LENGTH: 419
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-667-570A-3
```

```
  Query Match             86.9%;  Score 172;  DB 4;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 4.5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 13
US-10-182-263-1
; Sequence 1, Application US/10182263
; Patent No. 6630138
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-1

  Query Match             86.9%;  Score 172;  DB 4;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 4.5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 14
PCT-US92-10242-1
; Sequence 1, Application PC/TUS9210242
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 North Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10242
;      FILING DATE:  19921118
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/793,989
;      FILING DATE:  18-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  SCR0472P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  419 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..157
;      OTHER INFORMATION:  /note= "Protein C Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  158..169
;      OTHER INFORMATION:  /note= "Protein C Activation
;      OTHER INFORMATION:  Peptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  170..419
;      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
PCT-US92-10242-1

  Query Match             86.9%;  Score 172;  DB 5;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 4.5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 15
US-08-756-506-2
; Sequence 2, Application US/08756506
; Patent No. 5905185
;  GENERAL INFORMATION:
;    APPLICANT:  Garner, Ian
;    APPLICANT:  Cottingham, Ian R.
;    APPLICANT:  Temperley, Simon M.
;    APPLICANT:  Foster, Donald C.
;    APPLICANT:  Sprecher, Cindy A.
;    APPLICANT:  Prunkard, Donna E.
;    TITLE OF INVENTION:  PROTEIN C PRODUCTION IN TRANSGENIC
;    TITLE OF INVENTION:  ANIMALS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/756,506
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sawislak, Deborah A
;      REGISTRATION NUMBER:  37,438
;      REFERENCE/DOCKET NUMBER:  95-28
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6672
;      TELEFAX:  206-442-6678
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  460 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-756-506-2

  Query Match             86.9%;  Score 172;  DB 2;  Length 460;
  Best Local Similarity   75.0%;  Pred. No. 5e-21;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  || ||::|||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


Search completed: March  1, 2004, 10:12:20
Job time : 17.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 10:01:28 ; Search time 28 Seconds
(without alignments)
331.812 Million cell updates/sec

Title: SEQ1-32GLU-33ASP
Perfect score: 198
Sequence: 1 ANSFLXXLRHSSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	176	88.9	419	14	US-10-182-263-3	Sequence 3, Appli
2	176	88.9	419	14	US-10-182-263-4	Sequence 4, Appli
3	176	88.9	419	14	US-10-182-263-5	Sequence 5, Appli
4	176	88.9	419	15	US-10-168-407-3	Sequence 3, Appli
5	176	88.9	419	15	US-10-168-407-4	Sequence 4, Appli
6	172	86.9	44	14	US-10-298-330-1	Sequence 1, Appli
7	172	86.9	419	10	US-09-978-917A-4	Sequence 4, Appli
8	172	86.9	419	14	US-10-182-263-1	Sequence 1, Appli
9	172	86.9	419	15	US-10-168-407-1	Sequence 1, Appli
10	172	86.9	461	10	US-09-978-917A-2	Sequence 2, Appli
11	172	86.9	461	14	US-10-182-263-2	Sequence 2, Appli
12	172	86.9	461	15	US-10-168-407-2	Sequence 2, Appli
13	168	84.8	419	14	US-10-182-263-6	Sequence 6, Appli
14	168	84.8	419	15	US-10-168-407-5	Sequence 5, Appli
15	168	84.8	419	15	US-10-168-407-6	Sequence 6, Appli
16	109	55.1	44	14	US-10-298-330-2	Sequence 2, Appli
17	102	51.5	139	15	US-10-360-101-232	Sequence 232, App
18	102	51.5	488	14	US-10-348-504-44	Sequence 44, Appl
19	102	51.5	488	14	US-10-407-123-27	Sequence 27, Appl
20	95	48.0	44	14	US-10-298-330-18	Sequence 18, Appl
21	92	46.5	406	10	US-09-782-587B-3	Sequence 3, Appli
22	92	46.5	406	15	US-10-383-898-1	Sequence 1, Appli
23	92	46.5	466	14	US-10-017-122-2	Sequence 2, Appli
24	92	46.5	466	15	US-10-375-741-14	Sequence 14, Appl
25	91	46.0	42	16	US-10-038-854-97	Sequence 97, Appl
26	89	44.9	44	14	US-10-298-330-3	Sequence 3, Appli
27	89	44.9	406	10	US-09-782-587B-1	Sequence 1, Appli
28	89	44.9	406	14	US-10-109-498-1	Sequence 1, Appli
29	89	44.9	406	14	US-10-255-032-1	Sequence 1, Appli
30	89	44.9	406	14	US-10-281-727-1	Sequence 1, Appli
31	89	44.9	406	15	US-10-386-898-7	Sequence 7, Appli
32	88	44.4	405	15	US-10-360-101-225	Sequence 225, App
33	83	41.9	44	14	US-10-298-330-4	Sequence 4, Appli
34	80	40.4	40	14	US-10-298-330-23	Sequence 23, Appl
35	78	39.4	40	14	US-10-298-330-22	Sequence 22, Appl
36	78	39.4	40	14	US-10-298-330-25	Sequence 25, Appl
37	78	39.4	96	10	US-09-759-130B-313	Sequence 313, App
38	78	39.4	96	14	US-10-189-123-43	Sequence 43, Appl
39	78	39.4	96	14	US-10-188-495-43	Sequence 43, Appl
40	78	39.4	209	10	US-09-759-130B-312	Sequence 312, App
41	78	39.4	209	14	US-10-189-123-42	Sequence 42, Appl
42	78	39.4	209	14	US-10-188-495-42	Sequence 42, Appl
43	78	39.4	226	10	US-09-759-130B-310	Sequence 310, App
44	78	39.4	226	14	US-10-189-123-40	Sequence 40, Appl
45	78	39.4	226	14	US-10-188-495-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-10-182-263-3
; Sequence 3, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-3

Query Match 88.9%; Score 176; DB 14; Length 419;
Best Local Similarity 77.3%; Pred. No. 5.6e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||| ||| || | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44
```

RESULT 2
US-10-182-263-4
; Sequence 4, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:

; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-4

Query Match 88.9%; Score 176; DB 14; Length 419;
Best Local Similarity 77.3%; Pred. No. 5.6e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 3
US-10-182-263-5
; Sequence 5, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-5

Query Match 88.9%; Score 176; DB 14; Length 419;
Best Local Similarity 77.3%; Pred. No. 5.6e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 4
US-10-168-407-3
; Sequence 3, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13610
; CURRENT APPLICATION NUMBER: US/10/168,407
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-407-3

Query Match 88.9%; Score 176; DB 15; Length 419;
Best Local Similarity 77.3%; Pred. No. 5.6e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 5
US-10-168-407-4
; Sequence 4, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13610
; CURRENT APPLICATION NUMBER: US/10/168,407
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-407-4

Query Match 88.9%; Score 176; DB 15; Length 419;
Best Local Similarity 77.3%; Pred. No. 5.6e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 6
US-10-298-330-1
; Sequence 1, Application US/10298330
; Publication No. US20030100506A1
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary L.
; TITLE OF INVENTION: Modified Vitamin K-Dependent
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 09531-127001
; CURRENT APPLICATION NUMBER: US/10/298,330
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/497,591
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/302,239
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 08/955,636
; PRIOR FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 6, 7, 14, 16, 19, 20, 25, 26, 29
; OTHER INFORMATION: Xaa = gamma carboxyglutamic or glutamic acid
US-10-298-330-1

Query Match 86.9%; Score 172; DB 14; Length 44;
Best Local Similarity 95.5%; Pred. No. 2.2e-21;
Matches 42; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||||||||||||||||||||||||||||::|||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44
```

RESULT 7
US-09-978-917A-4
; Sequence 4, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Aps; Maxygen Holdings
; TITLE OF INVENTION: Protein C or activated protein C-like molecules
; FILE REFERENCE: 0219us310 - protein C
; CURRENT APPLICATION NUMBER: US/09/978,917A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-917A-4

Query Match 86.9%; Score 172; DB 10; Length 419;
Best Local Similarity 75.0%; Pred. No. 2.6e-20;
Matches 33; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

RESULT 8
US-10-182-263-1
; Sequence 1, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-1

Query Match 86.9%; Score 172; DB 14; Length 419;
Best Local Similarity 75.0%; Pred. No. 2.6e-20;
Matches 33; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

RESULT 9
US-10-168-407-1
; Sequence 1, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13610
; CURRENT APPLICATION NUMBER: US/10/168,407
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-407-1

Query Match 86.9%; Score 172; DB 15; Length 419;
Best Local Similarity 75.0%; Pred. No. 2.6e-20;
Matches 33; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db          1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

RESULT 10
US-09-978-917A-2
; Sequence 2, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Aps; Maxygen Holdings
; TITLE OF INVENTION: Protein C or activated protein C-like molecules
; FILE REFERENCE: 0219us310 - protein C
; CURRENT APPLICATION NUMBER: US/09/978,917A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(42)
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (43)...(461)
US-09-978-917A-2

Query Match 86.9%; Score 172; DB 10; Length 461;
Best Local Similarity 75.0%; Pred. No. 2.9e-20;
Matches 33; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86
```

RESULT 11
US-10-182-263-2
; Sequence 2, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14

```
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-2

  Query Match             86.9%;  Score 172;  DB 14;  Length 461;
  Best Local Similarity   75.0%;  Pred. No. 2.9e-20;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 12
US-10-168-407-2
; Sequence 2, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-2

  Query Match             86.9%;  Score 172;  DB 15;  Length 461;
  Best Local Similarity   75.0%;  Pred. No. 2.9e-20;
  Matches   33;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||::||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 13
US-10-182-263-6
; Sequence 6, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-6

  Query Match             84.8%;  Score 168;  DB 14;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-19;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 14
US-10-168-407-5
; Sequence 5, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-5

  Query Match             84.8%;  Score 168;  DB 15;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-19;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 15
US-10-168-407-6
; Sequence 6, Application US/10168407
; Publication No. US20030207435A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13610
;  CURRENT APPLICATION NUMBER: US/10/168,407
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-168-407-6

  Query Match             84.8%;  Score 168;  DB 15;  Length 419;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-19;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


Search completed: March  1, 2004, 10:11:10
Job time : 30 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:54:37 ; Search time 50 Seconds
(without alignments)
248.642 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	44	2	AAY18300	Aay18300 Modified
2	179	90.9	419	4	AAB82677	Aab82677 Human pro
3	179	90.9	419	4	AAB82678	Aab82678 Human pro
4	179	90.9	419	4	AAE08630	Aae08630 Human pro
5	176	89.3	44	2	AAY18301	Aay18301 Modified
6	174	88.3	419	4	AAB82676	Aab82676 Human pro
7	174	88.3	419	4	AAB82675	Aab82675 Human pro
8	174	88.3	419	4	AAE08628	Aae08628 Human pro
9	174	88.3	419	4	AAE08627	Aae08627 Human pro
10	174	88.3	419	4	AAE08629	Aae08629 Human pro
11	173	87.8	44	2	AAY18298	Aay18298 Modified
12	170	86.3	44	2	AAY18299	Aay18299 Modified
13	168	85.3	44	2	AAY18297	Aay18297 Modified
14	168	85.3	44	2	AAY18307	Aay18307 Modified
15	168	85.3	45	5	ABB79949	Abb79949 Human pro
16	164	83.2	45	5	ABB79946	Abb79946 Human pro
17	164	83.2	45	7	ADB71159	Adb71159 Human pro
18	162	82.2	45	5	ABB79950	Abb79950 Human pro
19	160	81.2	44	2	AAY18309	Aay18309 Modified
20	160	81.2	44	2	AAY18303	Aay18303 Human pro
21	160	81.2	44	4	AAB36402	Aab36402 Human pro
22	160	81.2	44	7	ADD50094	Add50094 Human vit
23	160	81.2	45	2	AAW75710	Aaw75710 Partial h
24	160	81.2	45	5	ABB79947	Abb79947 Human pro
25	160	81.2	45	7	ADB71155	Adb71155 Human pro
26	160	81.2	415	3	AAY56803	Aay56803 Truncated
27	160	81.2	419	2	AAR35760	Aar35760 Protein C
28	160	81.2	419	2	AAW72753	Aaw72753 Primary s
29	160	81.2	419	4	AAB82673	Aab82673 Wild-type
30	160	81.2	419	4	AAB36896	Aab36896 Human pro
31	160	81.2	419	4	AAB36897	Aab36897 Human pro
32	160	81.2	419	4	AAB36898	Aab36898 Human pro
33	160	81.2	419	4	AAB36894	Aab36894 Human pro
34	160	81.2	419	4	AAE08625	Aae08625 Human mat
35	160	81.2	419	5	AAU99005	Aau99005 Human Pro
36	160	81.2	419	5	AAU99006	Aau99006 Human Pro
37	160	81.2	419	5	AAU99008	Aau99008 Human Pro
38	160	81.2	419	5	AAU99018	Aau99018 Human Pro
39	160	81.2	419	5	AAU99026	Aau99026 Human Pro
40	160	81.2	419	5	AAU99037	Aau99037 Human Pro
41	160	81.2	419	5	AAU99049	Aau99049 Human Pro
42	160	81.2	419	5	AAU99063	Aau99063 Human Pro
43	160	81.2	419	5	AAU99072	Aau99072 Human Pro
44	160	81.2	419	5	AAU99083	Aau99083 Human Pro
45	160	81.2	419	5	AAU99084	Aau99084 Human Pro

ALIGNMENTS

RESULT 1
AAY18300
ID AAY18300 standard; peptide; 44 AA.
XX
AC AAY18300;
XX
DT 17-AUG-1999 (first entry)
XX
DE Modified GLA domain of vitamin K-dependent protein.
XX
KW GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .44
FT /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT acid"
XX
PN WO9920767-A1.
XX
PD 29-APR-1999.
XX
PF 20-OCT-1998; 98WO-US022152.
XX
PR 23-OCT-1997; 97US-00955636.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Nelsestuen GL;
XX
DR WPI; 1999-288309/24.
XX
PT Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT domain, useful for treating clotting disorders.
XX
PS Claim 9; Page 79; 86pp; English.
XX
CC This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC domain. The invention relates to a vitamin K-dependent polypeptide
CC comprising a modified GLA domain containing an amino acid substitution
CC which enhances membrane binding of the modified polypeptide as compared
CC to the native polypeptide. The polypeptide is used to treat a clotting
CC disorder by decreasing or increasing clot formation. Modification of the
CC GLA domain results in a protein which has enhanced membrane binding
CC affinity as compared to the native protein

```
XX
SQ   Sequence 44 AA;

  Query Match             90.9%;  Score 179;  DB 2;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-22;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||||||||||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 2
AAB82677
ID   AAB82677 standard; protein; 419 AA.
XX
AC   AAB82677;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (H10Q/S11G/Q32E/N33D/L194S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 10
FT                   /note= "His in wild-type protein"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26365.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 5; Page 54-55; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which His at position 10 of the mature wild-type protein C sequence (see
CC   AAB82673) is substituted with Gln, Ser at position 11 with Gly, Gln at
CC   position 32 with Glu, Asn at position 33 with Asp, and Leu at position
CC   194 with Ser. It is an example of protein C derivatives of the invention
CC   that have at least 2 amino acid substitutions, but which have increased
CC   anticoagulant activity and resistance to inactivation by serpins compared
CC   with the wild-type protein, while retaining the biological activity of
CC   the wild-type protein. A method of producing the derivatives using
CC   recombinant DNA methods is claimed. The protein C derivatives are useful
CC   for treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina), vascular
CC   occlusive disorders and hypercoagulable states, sepsis (in combination
CC   with bactericidal permeability increasing protein or with tissue factor
CC   pathway inhibitor), thrombotic disorders (in combination with an anti-
CC   platelet agent or by local delivery through an intracoronary catheter),
CC   protein C deficiency, acute arterial thrombotic occlusion,
CC   thromboembolism, or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts. Human patients with genetically predisposed
CC   prothrombotic disorders may be treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             90.9%;  Score 179;  DB 4;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 2.5e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 3
AAB82678
ID   AAB82678 standard; protein; 419 AA.
XX
AC   AAB82678;
```

```
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (H10Q/S11G/Q32E/N33D/L194S/T254S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 10
FT                   /note= "His in wild-type protein"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Misc-difference 254
FT                   /note= "Thr in wild-type protein"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 6; Page 56-57; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which His at position 10 of the wild-type protein C sequence (see
CC   AAB82673) is substituted with Gln, Ser at position 11 with Gly, Gln at
CC   position 32 with Glu, Asn at position 33 with Asp, Leu at position 194
CC   with Ser, and Thr at position 254 with Ser. It is an example of protein C
CC   derivatives of the invention that have at least 2 amino acid
CC   substitutions, but which have increased anticoagulant activity and
CC   resistance to inactivation by serpins compared with the wild-type
CC   protein, while retaining the biological activity of the wild-type
CC   protein. A method of producing the derivatives using recombinant DNA
CC   methods is claimed. The protein C derivatives are useful for treating
CC   coronary syndromes and disease states predisposing to thrombosis (e.g.
CC   myocardial infarction and unstable angina), vascular occlusive disorders
CC   and hypercoagulable states, sepsis (in combination with bactericidal
CC   permeability increasing protein or with tissue factor pathway inhibitor),
CC   thrombotic disorders (in combination with an anti-platelet agent or by
CC   local delivery through an intracoronary catheter), protein C deficiency,
CC   acute arterial thrombotic occlusion, thromboembolism, or stenosis in
CC   coronary, cerebral or peripheral arteries or in vascular grafts. Human
CC   patients with genetically predisposed prothrombotic disorders may be
CC   treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             90.9%;  Score 179;  DB 4;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 2.5e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 4
AAE08630
ID   AAE08630 standard; protein; 419 AA.
XX
AC   AAE08630;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #4.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US001221.
XX
PR   11-FEB-2000; 2000US-0181948P.
PR   14-MAR-2000; 2000US-0189199P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15228.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions.
XX
PS   Claim 6; Page 50-51; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and increased
CC   sensitivity to thrombin activation compared to wild type protein C, and
CC   retains the biological activity of the wild type human protein C. Protein
CC   C derivatives are useful in the manufacture of a medicament for the
CC   treatment of acute coronary syndromes e.g. myocardial infarction and
CC   unstable angina; and disease states predisposing to thrombosis; vascular
CC   occlusive disorders and hypercoagulable states e.g. disseminated
CC   intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC   sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC   syndrome; sepsis in combination with bacterial permeability increasing
CC   protein; thrombotic disorders in combination with an anti-platelet agent;
CC   protein C deficiency; acute arterial thrombotic occlusion,
CC   thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC   acid molecules of the invention are useful for treating humans with
CC   genetically predisposed prothrombotic disorders by gene therapy. The
CC   present sequence is human protein C derivative
XX
SQ   Sequence 419 AA;

  Query Match             90.9%;  Score 179;  DB 4;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 2.5e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||||  |  ||  ||||  ||  ||||||||||||||
Db          1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 5
AAY18301
ID   AAY18301 standard; peptide; 44 AA.
XX
AC   AAY18301;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 9; Page 82; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             89.3%;  Score 176;  DB 2;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 7.2e-22;
  Matches   43;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||:|||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLREGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 6
AAB82676
ID   AAB82676 standard; protein; 419 AA.
XX
AC   AAB82676;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S/T254S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
```

```
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Misc-difference 254
FT                   /note= "Thr in wild-type protein"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26364.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 4; Page 53-54; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which Ser at position 11 of the mature wild-type protein C sequence (see
CC   AAB82673) is substituted with Gly, Gln at position 32 with Glu, Asn at
CC   position 33 with Asp, Leu at position 194 with Ser, and Thr at position
CC   254 with Ser. It is an example of protein C derivatives of the invention
CC   that have at least 2 amino acid substitutions, but which have increased
CC   anticoagulant activity and resistance to inactivation by serpins compared
CC   with the wild-type protein, while retaining the biological activity of
CC   the wild-type protein. A method of producing the derivatives using
CC   recombinant DNA methods is claimed. The protein C derivatives are useful
CC   for treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina), vascular
CC   occlusive disorders and hypercoagulable states, sepsis (in combination
CC   with bactericidal permeability increasing protein or with tissue factor
CC   pathway inhibitor), thrombotic disorders (in combination with an anti-
CC   platelet agent or by local delivery through an intracoronary catheter),
CC   protein C deficiency, acute arterial thrombotic occlusion,
CC   thromboembolism, or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts. Human patients with genetically predisposed
CC   prothrombotic disorders may be treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             88.3%;  Score 174;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.7e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||| | ||  ||||  || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 7
AAB82675
ID   AAB82675 standard; protein; 419 AA.
XX
AC   AAB82675;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder; hypercoagulation;
KW   sepsis; protein C deficiency; occlusion; thromboembolism; stenosis;
KW   antibacterial; immunosuppressive; thrombolytic; cardiant; antianginal;
KW   anticoagulant; therapy; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .45
FT                   /note= "Gla domain"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Disulfide-bond  50. .69
FT   Disulfide-bond  59. .64
FT   Disulfide-bond  80. .89
FT   Disulfide-bond  98. .109
```

```
FT   Disulfide-bond  120. .133
FT   Disulfide-bond  141. .277
FT   Cleavage-site   156. .157
FT                   /note= "cleavage makes a 2-chain inactive precursor (155-
FT                   amino acid light chain attached via a disulfide bond to a
FT                   262-amino acid heavy chain)"
FT   Peptide         158. .169
FT                   /note= "activation peptide; removal activates the 2-chain
FT                   zymogen"
FT   Cleavage-site   169. .170
FT                   /note= "thrombin cleavage site"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Disulfide-bond  196. .212
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
FT   Disulfide-bond  331. .345
FT   Disulfide-bond  356. .384
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US000020.
XX
PR   02-FEB-2000; 2000US-0179801P.
PR   14-MAR-2000; 2000US-0189197P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26363.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute arterial
PT   thrombotic occlusion, and thromboembolism.
XX
PS   Claim 3; Page 52-53; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative in
CC   which Ser at amino acid position 11 of the mature wild-type protein C
CC   sequence (see AAB82673) is substituted with Gly, Gln at position 32 with
CC   Glu, Asn at position 33 with Asp, and Leu at position 194 with Ser. The
CC   protein is an example of protein C derivatives of the invention that have
CC   at least 2 amino acid substitutions, but which have increased
CC   anticoagulant activity and resistance to inactivation by serpins compared
CC   with the wild-type protein, while retaining the biological activity of
CC   the wild-type protein. A method of producing the derivatives using
CC   recombinant DNA methods is claimed. The protein C derivatives are useful
CC   for treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina), vascular
CC   occlusive disorders and hypercoagulable states, sepsis (in combination
CC   with bactericidal permeability increasing protein or with tissue factor
CC   pathway inhibitor), thrombotic disorders (in combination with an anti-
CC   platelet agent or by local delivery through an intracoronary catheter),
CC   protein C deficiency, acute arterial thrombotic occlusion,
CC   thromboembolism, or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts. Human patients with genetically predisposed
CC   prothrombotic disorders may be treated by gene therapy (all claimed)
XX
SQ   Sequence 419 AA;

  Query Match             88.3%;  Score 174;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.7e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44

RESULT 8
AAE08628
ID   AAE08628 standard; protein; 419 AA.
XX
AC   AAE08628;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #2.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US001221.
XX
PR   11-FEB-2000; 2000US-0181948P.
PR   14-MAR-2000; 2000US-0189199P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15226.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions.
XX
PS   Claim 4; Page 47-48; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and increased
CC   sensitivity to thrombin activation compared to wild type protein C, and
CC   retains the biological activity of the wild type human protein C. Protein
CC   C derivatives are useful in the manufacture of a medicament for the
CC   treatment of acute coronary syndromes e.g. myocardial infarction and
CC   unstable angina; and disease states predisposing to thrombosis; vascular
CC   occlusive disorders and hypercoagulable states e.g. disseminated
CC   intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC   sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC   syndrome; sepsis in combination with bacterial permeability increasing
CC   protein; thrombotic disorders in combination with an anti-platelet agent;
CC   protein C deficiency; acute arterial thrombotic occlusion,
CC   thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC   acid molecules of the invention are useful for treating humans with.
CC   genetically predisposed prothrombotic disorders by gene therapy. The
CC   present sequence is human protein C derivative
XX
SQ   Sequence 419 AA;

  Query Match             88.3%;  Score 174;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.7e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
                 ||||| || ||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44
```

```
RESULT 9
AAE08627
ID   AAE08627 standard; protein; 419 AA.
XX
AC   AAE08627;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #1.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US001221.
XX
PR   11-FEB-2000; 2000US-0181948P.
PR   14-MAR-2000; 2000US-0189199P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15225.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions.
XX
PS   Claim 3; Page 46-47; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and increased
CC   sensitivity to thrombin activation compared to wild type protein C, and
CC   retains the biological activity of the wild type human protein C. Protein
CC   C derivatives are useful in the manufacture of a medicament for the
CC   treatment of acute coronary syndromes e.g. myocardial infarction and
CC   unstable angina; and disease states predisposing to thrombosis; vascular
CC   occlusive disorders and hypercoagulable states e.g. disseminated
CC   intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC   sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC   syndrome; sepsis in combination with bacterial permeability increasing
CC   protein; thrombotic disorders in combination with an anti-platelet agent;
CC   protein C deficiency; acute arterial thrombotic occlusion,
CC   thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC   acid molecules of the invention are useful for treating humans with
CC   genetically predisposed prothrombotic disorders by gene therapy. The
CC   present sequence is human protein C derivative
XX
SQ   Sequence 419 AA;

  Query Match             88.3%;  Score 174;  DB 4;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1.7e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
                 ||||| || ||| | ||  |||| || ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44
```

```
RESULT 10
AAE08629
ID   AAE08629 standard; protein; 419 AA.
XX
AC   AAE08629;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #3.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "Encoded by CAA"
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US001221.
XX
PR   11-FEB-2000; 2000US-0181948P.
PR   14-MAR-2000; 2000US-0189199P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15227.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions.
XX
PS   Claim 5; Page 48-49; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and increased
CC   sensitivity to thrombin activation compared to wild type protein C, and
CC   retains the biological activity of the wild type human protein C. Protein
CC   C derivatives are useful in the manufacture of a medicament for the
CC   treatment of acute coronary syndromes e.g. myocardial infarction and
CC   unstable angina; and disease states predisposing to thrombosis; vascular
CC   occlusive disorders and hypercoagulable states e.g. disseminated
CC   intravascular coagulation (DIC), burns, transplantations, thalassaemia,
CC   sickle cell disease, viral haemorrhagic fever and haemolytic uremic
CC   syndrome; sepsis in combination with bacterial permeability increasing
CC   protein; thrombotic disorders in combination with an anti-platelet agent;
CC   protein C deficiency; acute arterial thrombotic occlusion,
CC   thromboembolism or stenosis in coronary, cerebral or peripheral arteries
CC   or in vascular grafts in combination with a thrombolytic agent. Nucleic
CC   acid molecules of the invention are useful for treating humans with
CC   genetically predisposed prothrombotic disorders by gene therapy. The
CC   present sequence is human protein C derivative
XX
SQ   Sequence 419 AA;

  Query Match             88.3%;  Score 174;  DB 4;  Length 419;
```

```
  Best Local Similarity   77.3%;  Pred. No. 1.7e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  |||  |  ||  ||||  ||  ||||||||||||||||
Db          1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 11
AAY18298
ID   AAY18298 standard; peptide; 44 AA.
XX
AC   AAY18298;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 7; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             87.8%;  Score 173;  DB 2;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 2.3e-21;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||| ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 12
AAY18299
ID   AAY18299 standard; peptide; 44 AA.
XX
AC   AAY18299;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US022152.
XX
PR   23-OCT-1997;   97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 8; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             86.3%;  Score 170;  DB 2;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 7.4e-21;
  Matches   42;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||: |||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 13
AAY18297
ID   AAY18297 standard; peptide; 44 AA.
XX
AC   AAY18297;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
```

```
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;    98WO-US022152.
XX
PR   23-OCT-1997;    97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Claim 6; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             85.3%;  Score 168;  DB 2;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 1.6e-20;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 14
AAY18307
ID   AAY18307 standard; peptide; 44 AA.
XX
AC   AAY18307;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                   acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;    98WO-US022152.
XX
PR   23-OCT-1997;    97US-00955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic acid
PT   domain, useful for treating clotting disorders.
XX
PS   Disclosure; Page 79; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein
XX
SQ   Sequence 44 AA;

  Query Match             85.3%;  Score 168;  DB 2;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 1.6e-20;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||||||  ||||||||||||||||||||||||||||||||||
Db          1 ANSFLXXLRDSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 15
ABB79949
ID   ABB79949 standard; protein; 45 AA.
XX
AC   ABB79949;
XX
DT   12-DEC-2002  (first entry)
XX
DE   Human protein C mutated Gla domain QGN.
XX
KW   Protein C; Gla domain; human; blood clotting; anticoagulant;
KW   thrombolytic; antiarteriosclerotic; cardiant; antiaggregant; mutant;
KW   mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild-type His substituted by Gln"
FT   Misc-difference 11
FT                   /note= "wild-type Ser substituted by Gly"
FT   Misc-difference 12
FT                   /note= "wild-type Ser substituted by Gln"
XX
PN   WO200270681-A1.
XX
PD   12-SEP-2002.
XX
PF   01-MAR-2002; 2002WO-SE000363.
XX
PR   02-MAR-2001; 2001US-0272466P.
XX
PA   (TACT-) TAC THROMBOSIS & COAGULATION AB.
XX
PI   Dahlbaeck B;
XX
DR   WPI; 2002-713449/77.
XX
PT   New variant blood coagulation component, useful for manufacturing a
PT   medicament for treating or preventing coagulation disorders, e.g.
PT   thrombosis, comprises an anticoagulant activity in the protein C-
PT   anticoagulant system of blood.
XX
```

PS Example 1; Page; 58pp; English.
XX
CC The present sequence is the protein sequence of a mutated Gla domain (N-
CC terminal amino acids 1-45) of human protein C. The mutated Gla domain
CC contains the substitution mutations H10Q, S11G and S12N. Protein C and
CC activated protein C variants comprising a mutated Gla domain are provided
CC by the invention. The variants contain at least 6, and optionally 7-10,
CC amino acid substitutions. A preferred mutant (designated QGNSEDY, see
CC ABB79946) has the mutations H10Q, S11G, S12N, D23S, Q32E, N33D and H44Y,
CC and shows greatly enhanced anticoagulant activity in standard in vitro
CC coagulation assays. The present mutant (designated QGN) was produced in
CC an example from the invention as a step toward the production of the
CC QGNSEDY mutant Gla domain, and shows little, if any, improvement in
CC anticoagulant activity over wild-type activated protein C. The invention
CC provides methods for producing the variants based on DNA technology, and
CC with the use of the variants for the treatment of coagulation disorders
CC such as thrombosis or APC resistance, or in diagnostic test systems for
CC assaying components of the protein C-anticoagulant system (all claimed).
CC The variants may also be used in treating arteriosclerosis, myocardial
CC infarction, and disseminated intravascular coagulation. Note: The present
CC sequence is not shown in the specification but is derived from the human
CC wild-type Gla domain sequence given on page 7 of the specification (see
CC ABB79947)
XX
SQ Sequence 45 AA;

Query Match 85.3%; Score 168; DB 5; Length 45;
Best Local Similarity 72.7%; Pred. No. 1.6e-20;
Matches 32; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||||:| | ||  ||||  || ||::||||||||||||
Db          1 ANSFLEELRQGNLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

Search completed: March 1, 2004, 10:01:23
Job time : 50 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:59:33 ; Search time 13.5 Seconds
(without alignments)
313.513 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	160	81.2	461	1	KXHU	protein C (activat
2	140	71.1	461	1	JX0210	protein C (activat
3	139	70.6	461	1	S18994	protein C (activat
4	122	61.9	456	1	KXBO	protein C (activat
5	115	58.4	482	1	EXRT	coagulation factor
6	114	57.9	492	1	EXBO	coagulation factor
7	110	55.8	488	1	EXHU	coagulation factor
8	101	51.3	443	2	I46932	coagulation factor
9	99	50.3	466	1	KFHU7	coagulation factor
10	86.5	43.9	617	2	S10511	thrombin (EC 3.4.2
11	86.5	43.9	618	2	A35827	thrombin (EC 3.4.2
12	86	43.7	475	1	EXCH	coagulation factor
13	85	43.1	407	1	KFBO7	coagulation factor
14	85	43.1	642	2	S53434	plasma protein S p
15	85	43.1	676	1	KXHUS	plasma protein S p
16	84	42.6	622	1	TBHU	thrombin (EC 3.4.2
17	81	41.1	646	2	S38819	plasma protein S -
18	80	40.6	452	1	A30351	coagulation factor
19	80	40.6	459	2	JQ0419	coagulation factor
20	80	40.6	461	1	KFHU	coagulation factor
21	80	40.6	675	1	KXBOS	plasma protein S p
22	78	39.6	642	2	S53433	plasma protein S p
23	78	39.6	675	1	KXRTS	plasma protein S p
24	73	37.1	416	1	KFBO	coagulation factor
25	72	36.5	625	1	TBBO	thrombin (EC 3.4.2
26	71	36.0	675	1	KXMSS	plasma protein S p
27	69.5	35.3	396	1	KXBOZ	plasma protein Z -
28	65.5	33.2	422	1	KXHUZ	plasma protein Z p
29	65	33.0	673	2	A48089	growth arrest-spec
30	64	32.5	674	2	I55476	growth potentiatin
31	63	32.0	678	2	B48089	growth arrest-spec
32	56.5	28.7	594	2	D84859	probable MAP kinas
33	54.5	27.7	603	2	C96575	probable MAP kinas
34	53.5	27.2	576	2	G96763	probable MAP kinas
35	53	26.9	606	2	T40556	hypothetical prote
36	50	25.4	1684	2	T02367	hypothetical prote
37	49	24.9	1363	2	I58375	protein-tyrosine k
38	48.5	24.6	323	2	T25948	hypothetical prote
39	48.5	24.6	510	2	E82918	ammonium transport
40	48	24.4	422	2	T39306	mitogen-activated
41	48	24.4	1235	2	D32433	VSG expression sit
42	48	24.4	1298	2	A48999	protein-tyrosine k
43	47.5	24.1	1089	1	S33727	platelet-derived g
44	47	23.9	182	2	JC1189	tyrosine kinase re
45	47	23.9	245	1	NDECR5	type II site-speci

ALIGNMENTS

RESULT 1
KXHU
protein C (activated) (EC 3.4.21.69) precursor - human
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A22331; A25426; A21781; A23789; A00927
R;Foster, D.C.; Yoshitake, S.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985
A;Title: The nucleotide sequence of the gene for human protein C.
A;Reference number: A22331; MUID:85270390; PMID:2991887
A;Accession: A22331
A;Molecule type: DNA
A;Residues: 1-461 <FOS1>
A;Cross-references: GB:M11228; NID:g190333; PIDN:AAA60166.1; PID:g190334
R;Plutzky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986
A;Title: Evolution and organization of the human protein C gene.
A;Reference number: A25426; MUID:86120978; PMID:3511471
A;Accession: A25426
A;Molecule type: DNA
A;Residues: 1-445,'L',446-461 <PLU>
A;Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332
R;Foster, D.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984
A;Title: Characterization of a cDNA coding for human protein C.
A;Reference number: A21781; MUID:84272714; PMID:6589623
A;Accession: A21781
A;Molecule type: mRNA
A;Residues: 'Q',107-461 <FOS2>
A;Cross-references: GB:K02059; NID:g190322; PIDN:AAA60164.1; PID:g190323
R;Beckmann, R.J.; Schmidt, R.J.; Santerre, R.F.; Plutzky, J.; Crabtree, G.R.; Long, G.L.
Nucleic Acids Res. 13, 5233-5247, 1985
A;Title! The structure and evolution of a 461 amino acid human protein C precursor and i
A;Reference number: A23789; MUID:85269639; PMID:2991859
A;Accession: A23789
A;Molecule type: mRNA
A;Residues: 1-461 <BEC>
A;Cross-references: GB:X02750; NID:g35689; PIDN:CAA26528.1; PID:g763120
R;Miletich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 265, 11397-11404, 1990
A;Title: Beta protein C is not glycosylated at asparagine 329. The rate of translation m
A;Reference number: A44605; MUID:90293094; PMID:1694179
A;Contents: annotation; carbohydrate binding sites; activation peptide
A;Note: the alpha form of protein C is glycosylated at Asn-329, and the beta form is not
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of facto
A;Reference number: A44606; MUID:92184750; PMID:1544894
A;Contents: annotation; beta-hydroxyaspartic acid
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that in
ivation of factor Va is strongly enhanced by complexing with protein S. Protein C also f

```
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction,
C;Genetics:
A;Gene: GDB:PROC
A;Cross-references: GDB:120317; OMIM:176860
A;Map position: 2q13-2q21
A;Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-86/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-197/Product: protein C light chain #status predicted <LCH>
F;92-131/Domain: EGF homology <EG1>
F;140-175/Domain: EGF homology <EG2>
F;200-461/Product: protein C heavy chain #status predicted <HCH>
F;200-211/Domain: activation peptide #status experimental <APT>
F;212-445/Domain: trypsin homology <TRY>
F;48,49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp
F;59-64,92-105,101-120,122-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/D
F;106-111/Disulfide bonds: #status predicted
F;110/Binding site: carbohydrate (Thr) (covalent) #status absent
F;113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;139,290,355/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;211-212/Cleavage site: Arg-Leu (thrombin) #status experimental
F;253,299,402/Active site: His, Asp, Ser #status predicted
F;371/Binding site: carbohydrate (Asn) (covalent) (partial) #status atypical

  Query Match             81.2%;  Score 160;  DB 1;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 8.6e-18;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  || | ||  ||||  || ||::||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
JX0210
protein C (activated) (EC 3.4.21.69) precursor - mouse
N;Alternate names: vitamin K-dependent serine proteinase
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: JX0210
R;Tada, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.
J. Biochem. 111, 491-495, 1992
A;Title: Isolation and characterization of a mouse protein C cDNA.
A;Reference number: JX0210; MUID:92316897; PMID:1618739
A;Accession: JX0210
A;Molecule type: mRNA
A;Residues: 1-461 <TAD>
A;Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386
A;Experimental source: liver
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-33/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;34-41/Domain: propeptide #status predicted <PRO>
F;42-196,199-461/Product: protein C #status predicted <PRC>
F;42-196/Domain: light chain #status predicted <PCL>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;199-461/Domain: heavy chain #status predicted <PCH>
F;199-211/Domain: activation peptide #status predicted <ACT>
F;212-461/Product: vitamin K-dependent serine proteinase #status predicted <VIT>
F;212-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-319,238-254,373-387,398-426/Disulfide bonds: #stat
F;214,290,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;253,299,402/Active site: His, Asp, Ser #status predicted

  Query Match             71.1%;  Score 140;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 1.5e-14;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  |: ||::|:||||||  |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
S18994
protein C (activated) (EC 3.4.21.69) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C;Accession: S18994; S24312
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
submitted to the EMBL Data Library, February 1992
A;Description: The cDNA cloninig and mRNA expression of rat protein C.
A;Reference number: S18994
A;Accession: S18994
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
Biochim. Biophys. Acta 1131, 329-332, 1992
A;Title: The cDNA cloning and mRNA expression of rat protein C.
A;Reference number: S24312; MUID:92329550; PMID:1627650
A;Accession: S24312
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA2>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; glycoprotein; hydrolase; serine proteinase
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-461/Product: protein C #status predicted <PRC>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;213-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-320,239-255,373-387,398-426/Disulfide bonds: #stat
F;215,291,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;254,300,402/Active site: His, Asp, Ser #status predicted

  Query Match             70.6%;  Score 139;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 2.1e-14;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  |: ||::|:||||||  |:
Db         42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
KXBO
protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Nov-1980 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A26250; A18385; A18386; A00928
R;Long, G.L.; Balagaje, R.M.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984
A;Title: Cloning and sequence of liver cDNA coding for bovine protein C.
A;Reference number: A26250; MUID:85014826; PMID:6091100
A;Accession: A26250
A;Molecule type: mRNA
A;Residues: 1-456 <LON>
R;Fernlund, P.; Stenflo, J.
```

```
J. Biol. Chem. 257, 12170-12179, 1982
A;Title: Amino acid sequence of the light chain of bovine protein C.
A;Reference number: A18385; MUID:83007325; PMID:6896876
A;Accession: A18385
A;Molecule type: protein
A;Residues: 40-194 <PER>
A;Note: 82-Lys was also found
R;Drakenberg, T.; Fernlund, P.; Roepstorff, P.; Stenflo, J.
Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983
A;Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.
A;Reference number: A19316; MUID:83169769; PMID:6572939
A;Contents: annotation; revision to residue 110
R;Stenflo, J.; Fernlund, P.
J. Biol. Chem. 257, 12180-12190, 1982
A;Title: Amino acid sequence of the heavy chain of bovine protein C.
A;Reference number: A18386; MUID:83007326; PMID:6896877
A;Accession: A18386
A;Molecule type: protein
A;Residues: 197-454,'PV' <STE>
R;Esmon, N.L.; DeBault, L.E.; Esmon, C.T.
J. Biol. Chem. 258, 5548-5553, 1983
A;Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless p
A;Reference number: A37541; MUID:83213513; PMID:6304092
A;Contents: annotation; activation; calcium binding
R;Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.
J. Biol. Chem. 258, 5554-5560, 1983
A;Title: Structural changes required for activation of protein C are induced by Ca2+ bin
A;Reference number: A37542; MUID:83213514; PMID:6406503
A;Contents: annotation; activation; calcium binding
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this reacti
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with strc
cognition of the thrombin-thrombomodulin complex.
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>
F;24-83/Domain: Gla domain homology <GLA>
F;30-39/Domain: propeptide #status predicted <PRO>
F;40-194/Product: protein C light chain #status experimental <LCH>
F;98-128/Domain: EGF homology <EG1>
F;137-172/Domain: EGF homology <EG2>
F;197-456/Product: protein C heavy chain #status experimental <HCH>
F;197-210/Domain: activation peptide #status experimental <APT>
F;211-440/Domain: trypsin homology <TRY>
F;45,46,53,55,58,59,62,64,65,68,74/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;119-128,137-148,144-157,159-172,180-318,237-253,368-382,393-421/Disulfide bonds: #stat
F;136,289,350/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;252,298,397/Active site: His, Asp, Ser #status predicted
F;366/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             61.9%;  Score 122;  DB 1;  Length 456;
  Best Local Similarity   50.0%;  Pred. No. 1.2e-11;
  Matches   21;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  || |:: | |   :|:|  |: ||:: :||:||||
Db         40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 5
EXRT
coagulation factor Xa (EC 3.4.21.6) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Jan-1995 #sequence_revision 07-Feb-1997 #text_change 08-Dec-2000
C;Accession: S49075; JC4670; PS0191; PS0190; I62745
R;Stanton, C.; Ross, P.; Hutson, S.; Wallin, R.
Thromb. Res. 80, 63-73, 1995
A;Title: Evidence for competition between vitamin K-dependent clotting factors for intra
A;Reference number: A58498; MUID:96093366; PMID:8578539
A;Accession: S49075
A;Molecule type: mRNA
A;Residues: 1-482 <STA1>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Note: submitted to the EMBL Data Library, June 1994
A;Note: neither the complete nucleic acid sequence nor the complete translation are show
R;Stanton, C.; Ross, R.P.; Hutson, S.; Wallin, R.
Gene 169, 269-273, 1996
A;Title: Processing and expression of rat and human clotting factor-X-encoding cDNAs.
A;Reference number: JC4670; MUID:96194815; PMID:8647460
A;Accession: JC4670
A;Molecule type: mRNA
A;Residues: 1-482 <STA2>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Experimental source: Cos-1 cell
R;Enjyoji, K.; Miyazaki, K.; Kato, H.
J. Biochem. 109, 890-898, 1991
A;Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat pla
A;Reference number: PS0190; MUID:92041742; PMID:1718949
A;Accession: PS0191
A;Molecule type: protein
A;Residues: 41-58,'X',60-65 <ENJ1>
A;Accession: PS0190
A;Molecule type: protein
A;Residues: 183-186,'X',188-207 <ENJ2>
R;Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A;Title: Analysis of the partial nucleotide sequences and deduced primary structures of t
A;Reference number: I46196; MUID:94222160; PMID:8168596
A;Accession: I62745
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 295-383,'G',385-455 <MUR>
A;Cross-references: GB:D21215; NID:g415309; PIDN:BAA04756.1; PID:g455396
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status predicted <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-482/Product: coagulation factor X heavy chain #status predicted <HCH>
F;183-231/Domain: activation peptide #status predicted <APT>
F;232-482/Product: coagulation factor Xa heavy chain #status predicted <ACT>
F;232-460/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-340,238-243,259-275,388-402,41
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;187/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;208/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;231-232/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat
F;274,320,417/Active site: His, Asp, Ser #status predicted

  Query Match             58.4%;  Score 115;  DB 1;  Length 482;
  Best Local Similarity   43.2%;  Pred. No. 1.6e-10;
  Matches   19;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||||   :::|:| | |:  || |  |: :||| : |  ||:|:
Db         41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 6
EXBO
coagulation factor Xa (EC 3.4.21.6) precursor - bovine
N;Alternate names: Stuart factor
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
```

C;Accession: A22867; A14997; A12030; A34412; S39414; A00925
R;Fung, M.R.; Campbell, R.M.; MacGillivray, T.A.
Nucleic Acids Res. 12, 4481-4492, 1984
A;Title: Blood coagulation factor X mRNA encodes a single polypeptide chain containing a
A;Reference number: A22867; MUID:84247315; PMID:6330671
A;Accession: A22867
A;Molecule type: mRNA
A;Residues: 1-487 <FUN>
A;Cross-references: GB:X00673; NID:g192; PIDN:CAA25286.1; PID:g193
R;Enfield, D.L.; Ericsson, L.H.; Fujikawa, K.; Walsh, K.A.; Neurath, H.; Titani, K.
Biochemistry 19, 659-667, 1980
A;Title: Amino acid sequence of the light chain of bovine factor X-1 (Stuart factor).
A;Reference number: A14997; MUID:80130563; PMID:6766735
A;Accession: A14997
A;Molecule type: protein
A;Residues: 41-102,'N',104-180 <ENF>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood cc
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Contents: annotation; revision to residue 103
R;Titani, K.; Fujikawa, K.; Enfield, D.L.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.
Proc. Natl. Acad. Sci. U.S.A. 72, 3082-3086, 1975
A;Title: Bovine factor X-1 (Stuart factor): amino-acid sequence of heavy chain.
A;Reference number: A12030; MUID:76053069; PMID:1059093
A;Accession: A12030
A;Molecule type: protein
A;Residues: 183-292,294-295,'GDE',299-334,336-348,'AE',351-354,356-441,'GKFG',446-492 <T
A;Note: carbohydrate binding sites and disulfide bonds were determined
R;Persson, E.; Selander, M.; Linse, S.; Drakenberg, T.; Oehlin, A.K.; Stenflo, J.
J. Biol. Chem. 264, 16897-16904, 1989
A;Title: Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal
A;Reference number: A34412; MUID:89380326; PMID:2789221
A;Accession: A34412
A;Molecule type: protein
A;Residues: 85-126 <PER>
A;Note: beta-hydroxyaspartic acid site
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39414
A;Molecule type: protein
A;Residues: 183-196;199-209;216-233 <INO>
A;Note: carbohydrate binding sites
R;Titani, K.; Hermodson, M.A.; Fujikawa, K.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.; D
Biochemistry 11, 4899-4903, 1972
A;Title: Bovine factor X-1a (activated Stuart factor). Evidence of homology with mammali
A;Reference number: A12453; MUID:73053314; PMID:4264286
A;Contents: annotation; active site
R;Fujikawa, K.; Titani, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 72, 3359-3363, 1975
A;Title: Activation of bovine factor X (Stuart factor): conversion of factor Xaalpha to
A;Reference number: A13504; MUID:76053121; PMID:1059122
A;Contents: annotation; activation
R;Sugo, T.; Bjork, I.; Holmgren, A.; Stenflo, J.
J. Biol. Chem. 259, 5705-5710, 1984
A;Title: Calcium-binding properties of bovine factor X lacking the gamma-carboxyglutamic
A;Reference number: A38024; MUID:84185716; PMID:6546930
A;Contents: annotation; calcium binding
R;Morita, T.; Jackson, C.M.
J. Biol. Chem. 261, 4008-4014, 1986
A;Reference number: A38025; MUID:86140210; PMID:3949800
A;Contents: annotation; sulfate binding
C;Comment: Factor Xa converts prothrombin to thrombin during blood clotting.
C;Comment: The two chains are formed from a single-chain precursor by the excision of tw
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway), c
activation.
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with stro
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
C;Genetics:
A;Gene: F10
A;Map position: 13q34

C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-180/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-492/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-233/Domain: activation peptide #status experimental <APT>
F;234-492/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;234-461/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-341/Disulfide bonds: #status p
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;200/Binding site: sulfate (Tyr) (covalent) (partial) #status experimental
F;208,485/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;233-234/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat
F;240-245,260-276,389-403,414-442/Disulfide bonds: #status experimental
F;275,321,418/Active site: His, Asp, Ser #status predicted

Query Match 57.9%; Score 114; DB 1; Length 492;
Best Local Similarity 45.5%; Pred. No. 2.4e-10;
Matches 20; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ::||:| | |:   |    |: :||| : |  ||||:
Db         41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84
```

RESULT 7
EXHU
coagulation factor Xa (EC 3.4.21.6) precursor [validated] - human
N;Alternate names: Stuart factor
C;Species: Homo sapiens (man)
C;Date: 15-Nov-1984 #sequence_revision 02-May-1994 #text_change 08-Dec-2000
C;Accession: A24478; JQ0917; A42485; A25853; A22208; A21284; A20362; S39415; I54051; A00
R;Leytus, S.P.; Foster, D.C.; Kurachi, K.; Davie, E.W.
Biochemistry 25, 5098-5102, 1986
A;Title: Gene for human Factor X: a blood coagulation factor whose gene organization is
A;Reference number: A24478; MUID:87026600; PMID:3768336
A;Accession: A24478
A;Molecule type: DNA
A;Residues: 1-488 <LEY>
A;Cross-references: GB:L29433; GB:M14327; NID:g459809; PIDN:AAA52764.1; PID:g182831
R;Messier, T.L.; Pittman, D.D.; Long, G.L.; Kaufman, R.J.; Church, W.R.
Gene 99, 291-294, 1991
A;Title: Cloning and expression in COS-1 cells of a full-length cDNA encoding human coag
A;Reference number: JQ0917; MUID:91216473; PMID:1902434
A;Accession: JQ0917
A;Molecule type: mRNA
A;Residues: 1-488 <MES>
A;Cross-references: GB:M57285; NID:g182389; PIDN:AAA52421.1; PID:g182390
R;Miao, C.H.; Leytus, S.P.; Chung, D.W.; Davie, E.W.
J. Biol. Chem. 267, 7395-7401, 1992
A;Title: Liver-specific expression of the gene coding for human factor X, a blood coagul
A;Reference number: A42485; MUID:92218390; PMID:1313796
A;Accession: A42485
A;Molecule type: DNA
A;Residues: 1-15 <MIA>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:93780, NCBIP:93787)
R;Kaul, R.K.; Hildebrand, B.; Roberts, S.; Jagadeeswaran, P.
Gene 41, 311-314, 1986
A;Title: Isolation and characterization of human blood-coagulation factor X cDNA.
A;Reference number: A25853; MUID:86221713; PMID:3011603
A;Accession: A25853
A;Molecule type: mRNA
A;Residues: 19-284,'E',289-488 <KAU>

A;Cross-references: GB:M22613; NID:g180335; PIDN:AAA51984.1; PID:g180336
R;Fung, M.R.; Hay, C.W.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 3591-3595, 1985
A;Title: Characterization of an almost full-length cDNA coding for human blood coagulati
A;Reference number: A22208; MUID:85216545; PMID:2582420
A;Accession: A22208
A;Molecule type: mRNA
A;Residues: 13-441,'S',443-488 <FUN>
A;Cross-references: GB:K03194; NID:g182840; PIDN:AAA52490.1; PID:g182841
R;Leytus, S.P.; Chung, D.W.; Kisiel, W.; Kurachi, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 3699-3702, 1984
A;Title: Characterization of a cDNA coding for human factor X.
A;Reference number: A21284; MUID:84222026; PMID:6587384
A;Accession: A21284
A;Molecule type: mRNA
A;Residues: 13-284,'E',289-488 <LE2>
A;Cross-references: GB:K01886
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.; Sasagawa, T.; Howald, W.N.; Kwa, E.Y.; Weins
Biochemistry 22, 2875-2884, 1983
A;Title: Complete amino acid sequence of the light chain of human blood coagulation fact
A;Reference number: A20362; MUID:83257207; PMID:6871167
A;Accession: A20362
A;Molecule type: protein
A;Residues: 41-179 <MCM>
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39415
A;Molecule type: protein
A;Residues: 183-234 <INO>
A;Note: glycosylation sites
A;Note: identification and characterization of beta-hydroxyaspartic acid
R;Jagadeeswaran, P.; Reddy, S.V.; Rao, K.J.; Hamsabhushanam, K.; Lyman, G.
Gene 84, 517-519, 1989
A;Title: Cloning and characterization of the 5' end (exon 1) of the gene encoding human
A;Reference number: I54051; MUID:90128299; PMID:2612918
A;Accession: I54051
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-23 <RES>
A;Cross-references: GB:M33297; NID:g183860; PIDN:AAA52636.1; PID:g553330
R;Padmanabhan, K.; Padmanabhan, K.P.; Tulinsky, A.; Park, C.H.; Bode, W.; Huber, R.; Bla
J. Mol. Biol. 232, 947-966, 1993
A;Title: Structure of human des(1-45) factor Xa at 2.2 angstroms resolution.
A;Reference number: A49458; MUID:93360277; PMID:8355279
A;Contents: annotation; X-ray crystallography, 2.2 angstroms
C;Comment: The two chains held together by one disulfide bond are formed from a single-c
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway) or
C;Genetics:
A;Gene: GDB:F10
A;Cross-references: GDB:119890; OMIM:227600
A;Map position: 13q34-13q34
A;Introns: 24/1; 77/3; 86/1; 124/1; 150/3; 249/3; 289/1
A;Note: deficiency of this factor causes Stuart disease
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-488/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-234/Domain: activation peptide #status experimental <APT>
F;235-488/Product: coagulation factor Xa heavy chain #status experimental <ACT>
F;235-462/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;57-62/Disulfide bonds: #status predicted
F;90-101,95-110,112-121,129-140,136-149,151-164,172-342,241-246,261-277,390-404,415-443/

F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;199,211/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;221,231/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;234-235/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat
F;276,322,419/Active site: His, Asp, Ser #status experimental

Query Match 55.8%; Score 110; DB 1; Length 488;
Best Local Similarity 43.2%; Pred. No. 1.1e-09;
Matches 19; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
Db 41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84

RESULT 8
I46932
coagulation factor VII - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 12-Feb-1999
C;Accession: I46932
R;Brothers, A.B.; Clarke, B.J.; Sheffield, W.P.; Blajchman, M.A.
Thromb. Res. 69, 231-238, 1993
A;Title: Complete nucleotide sequence of the cDNA encoding rabbit coagulation factor VII
A;Reference number: I46932; MUID:93190306; PMID:8383365
A;Accession: I46932
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-443 <BRO>
A;Cross-references: GB:S56300; NID:g266294; PID:g266295
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
F;24-83/Domain: Gla domain homology <GLA>
F;89-120/Domain: EGF homology <EG1>
F;130-166/Domain: EGF homology <EG2>
F;192-425/Domain: trypsin homology <TRY>

Query Match 51.3%; Score 101; DB 2; Length 443;
Best Local Similarity 46.3%; Pred. No. 2.7e-08;
Matches 19; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
Db 40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80

RESULT 9
KFHU7
coagulation factor VIIa (EC 3.4.21.21) precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 19-May-1994 #text_change 08-Dec-2000
C;Accession: A28322; A23819; A31186; B31186; S63524
R;O'Hara, P.J.; Grant, F.J.; Haldeman, B.A.; Gray, C.L.; Insley, M.Y.; Hagen, F.S.; Murr
Proc. Natl. Acad. Sci. U.S.A. 84, 5158-5162, 1987
A;Title: Nucleotide sequence of the gene coding for human factor VII, a vitamin K-depend
A;Reference number: A28322; MUID:87260948; PMID:3037537
A;Accession: A28322
A;Molecule type: DNA
A;Residues: 1-466 <OHA>
A;Cross-references: GB:J02933; NID:g180333; PIDN:AAA51983.1; PID:g180334
R;Hagen, F.S.; Gray, C.L.; O'Hara, P.; Grant, F.J.; Saari, G.C.; Woodbury, R.G.; Hart, C
Proc. Natl. Acad. Sci. U.S.A. 83, 2412-2416, 1986
A;Title: Characterization of a cDNA coding for human factor VII.
A;Reference number: A23819; MUID:86205965; PMID:3486420
A;Accession: A23819
A;Molecule type: mRNA
A;Residues: 1-466 <HAG>
A;Cross-references: GB:M13232; NID:g182799; PIDN:AAA88040.1; PID:g182801
R;Thim, L.; Bjoern, S.; Christensen, M.; Nicolaisen, E.M.; Lund-Hansen, T.; Pedersen, A.
Biochemistry 27, 7785-7793, 1988
A;Title: Amino acid sequence and posttranslational modifications of human factor VII-a f
A;Reference number: A90539; MUID:89088153; PMID:3264725
A;Accession: A31186
A;Molecule type: protein

```
A;Residues: 61-212 <THI>
A;Accession: B31186
A;Molecule type: protein
A;Residues: 213-466 <TH2>
R;Bjoern, S.; Foster, D.C.; Thim, L.; Wiberg, F.C.; Christensen, M.; Komiyama, Y.; Peder
J. Biol. Chem. 266, 11051-11057, 1991
A;Title: Human plasma and recombinant factor VII. Characterization of O-glycosylations a
A;Reference number: A40529; MUID:91250411; PMID:1904059
A;Contents: annotation; carbohydrate binding sites
R;Persson, E.; Petersen, L.C.
Eur. J. Biochem. 234, 293-300, 1995
A;Title: Structurally and functionally distinct Ca(2+) binding sites in the gamma-carbox
A;Reference number: S63524; MUID:96096752; PMID:8529655
A;Accession: S63524
A;Molecule type: protein
A;Residues: 61-65;99-103;105-109;213-217;308-312 <PER>
C;Genetics:
A;Gene: GDB:F7
A;Cross-references: GDB:119897; OMIM:227500
A;Map position: 13q34-13q34
A;Introns: 22/1; 44/1; 97/3; 106/1; 144/1; 191/1; 227/3; 269/1
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the prese
oagulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-60/Domain: propeptide #status predicted <PRO>
F;45-104/Domain: Gla domain homology <GLA>
F;61-212/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;110-141/Domain: EGF homology <EG1>
F;151-187/Domain: EGF homology <EG2>
F;213-466/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;213-447/Domain: trypsin homology <TRY>
F;66,67,74,76,79,80,85,86,89,95/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;77-82,110-121,115-130,132-141,151-162,158-172,174-187,195-322,219-224,238-254,370-389,
F;112,120/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;123/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status absent
F;205,382/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;212-213/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;253,302,404/Active site: His, Asp, Ser #status predicted
F;350-351/Cleavage site: Arg-Gly (coagulation factor Xa) #status predicted

  Query Match             50.3%;  Score 99;  DB 1;  Length 466;
  Best Local Similarity   48.8%;  Pred. No. 6.1e-08;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||:||  || ||| | |    | |  |: ||:| : |  ||
Db         61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 10
S10511
thrombin (EC 3.4.21.5) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 03-May-2002
C;Accession: S10511; A60576; B42696
R;Dihanich, M.; Monard, D.
Nucleic Acids Res. 18, 4251, 1990
A;Title: cDNA sequence of rat prothrombin.
A;Reference number: S10511; MUID:90332426; PMID:2377469
A;Accession: S10511
A;Molecule type: mRNA
A;Residues: 1-617 <DIH>
A;Cross-references: EMBL:X52835; NID:g56969; PIDN:CAA37017.1; PID:g56970
R;Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.
Endocrinology 126, 167-175, 1990
A;Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
A;Reference number: A60576; MUID:90091942; PMID:2293980
A;Accession: A60576
A;Molecule type: protein
A;Residues: 44-58 <HEN>
A;Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat ute
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: B42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 383-617,'E' <BAN>
A;Cross-references: GB:M81397
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-617/Product: prothrombin #status experimental <PMAT>
F;109-187/Domain: kringle homology <KR1>
F;215-292/Domain: kringle homology <KR2>
F;360-609/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,5
F;402,458,564/Active site: His, Asp, Ser #status predicted

  Query Match             43.9%;  Score 86.5;  DB 2;  Length 617;
  Best Local Similarity   42.2%;  Pred. No. 8.4e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||:|:| | |:  | : |   |  || ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 11
A35827
thrombin (EC 3.4.21.5) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
C;Accession: A35827; A42696; S12081
R;Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.
DNA Cell Biol. 9, 487-498, 1990
A;Title: Characterization of the cDNA coding for mouse prothrombin and localization of t
A;Reference number: A35827; MUID:91025551; PMID:2222810
A;Accession: A35827
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-618 <DEG>
A;Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814
A;Experimental source: strain C57BL/6
A;Note: the data were obtained from females resulting from the cross of M. domesticus an
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: A42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 384-618,'E' <BAN>
A;Cross-references: GB:M81394
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-618/Product: prothrombin B #status predicted <MAT>
F;109-187/Domain: kringle homology <KR1>
F;215-293/Domain: kringle homology <KR2>
F;361-610/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,5
F;403,459,565/Active site: His, Asp, Ser #status predicted

  Query Match             43.9%;  Score 86.5;  DB 2;  Length 618;
```

```
  Best Local Similarity   42.2%;  Pred. No. 8.4e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||:|:| | |:   | :  |    |   ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 12
EXCH
coagulation factor Xa (EC 3.4.21.6) precursor - chicken
N;Alternate names: virus-activating proteinase
C;Species: Gallus gallus (chicken)
C;Date: 12-Feb-1993 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C;Accession: S15838; S20380; S20381
R;Suzuki, H.; Harada, A.; Hayashi, Y.; Wada, K.; Asaka, J.; Gotoh, B.; Ogasawara, T.; Na
FEBS Lett. 283, 281-285, 1991
A;Title: Primary structure of the virus activating protease from chick embryo. Its ident
A;Reference number: S15838; MUID:91257322; PMID:2044767
A;Accession: S15838
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-475 <SUZ>
A;Cross-references: DDBJ:D00844; NID:g222869; PIDN:BAA00724.1; PID:g222870
R;Gotoh, B.; Yamauchi, F.; Ogasawara, T.; Nagai, Y.
FEBS Lett. 296, 274-278, 1992
A;Title: Isolation of factor Xa from chick embryo as the amniotic endoprotease responsib
A;Reference number: S20380; MUID:92164779; PMID:1537403
A;Accession: S20380
A;Molecule type: protein
A;Residues: 41-55 <GO2>
A;Accession: S20381
A;Molecule type: protein
A;Residues: 241-246,'X',248-251,'X',253-261 <GOT>
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-185/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-167/Domain: EGF homology <EG2>
F;186-475/Product: coagulation factor X heavy chain #status predicted <HCH>
F;186-240/Domain: activation peptide #status predicted <APT>
F;241-475/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;241-468/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-152,154-167,175-348,247-252,267-283,396-410,42
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;196,207,228,285/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;282,328,425/Active site: His, Asp, Ser #status predicted

  Query Match             43.7%;  Score 86;  DB 1;  Length 475;
  Best Local Similarity   40.5%;  Pred. No. 7.8e-06;
  Matches   17;  Conservative    7;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  ::||:: | |    |    |:  ||| : |  ||:
Db         41 ANSFLEEMKQGNIERECNEERCSKEEAREAFEDNEKTEEFWN 82


RESULT 13
KFBO7
coagulation factor VIIa (EC 3.4.21.21) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-May-1990 #sequence_revision 23-Mar-1995 #text_change 16-Jul-1999
C;Accession: A31979; C20274
R;Takeya, H.; Kawabata, S.; Nakagawa, K.; Yamamichi, Y.; Miyata, T.; Iwanaga, S.
J. Biol. Chem. 263, 14868-14877, 1988
A;Title: Bovine factor VII. Its purification and complete amino acid sequence.
A;Reference number: A31979; MUID:89008362; PMID:3049594
A;Accession: A31979
A;Molecule type: protein
A;Residues: 1-407 <TAK>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Accession: C20274
A;Molecule type: protein
A;Residues: 58-62,'X',64-68 <MCM>
A;Note: the residue designated 'X' was determined to be hydroxyaspartic acid
R;Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanaga,
J. Biochem. 104, 867-868, 1988
A;Title: A new trisaccharide sugar chain linked to a serine residue in bovine blood coagu
A;Reference number: A44556; MUID:89213999; PMID:3149637
A;Contents: annotation
A;Note: structure and location of covalently bound carbohydrate
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the preser
gulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-152/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;1-44/Domain: Gla domain homology (fragment) <GLA>
F;50-81/Domain: EGF homology <EG1>
F;91-127/Domain: EGF homology <EG2>
F;153-407/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;153-387/Domain: trypsin homology <TRY>
F;6,7,14,16,19,20,25,26,29,34,35/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;17-22,50-61,55-70,72-81,91-102,98-112,114-127,135-262,159-164,178-194,310-329,340-368/I
F;52/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;63/Modified site: erythro-beta-hydroxyaspartic acid (Asp) (partial) #status experiment
F;145,203/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;152-153/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;193,242,344/Active site: His, Asp, Ser #status predicted
F;290-291/Cleavage site: Arg-Gly (coagulation factor Xa) #status experimental

  Query Match             43.1%;  Score 85;  DB 1;  Length 407;
  Best Local Similarity   43.9%;  Pred. No. 9.6e-06;
  Matches   18;  Conservative    3;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  |  ||| | |   :| |  |  || : :  |  ||
Db          1 ANGFLEELLPGSLERECREELCSFEEAHEIFRNEERTRQFW 41


RESULT 14
S53434
plasma protein S precursor, vitamin K dependent - rhesus macaque (fragment)
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 16-Jul-1999
C;Accession: S53434
R;Greengard, J.S.; Fernandez, J.A.; Radtke, K.P.; Griffin, J.H.
Biochem. J. 305, 397-403, 1995
A;Title: Identification of candidate residues for interaction of protein S with C4b bind
A;Reference number: S53433; MUID:95134217; PMID:7832752
A;Accession: S53434
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-642 <GRE>
A;Cross-references: EMBL:L31380
A;Experimental source: tissue type liver
A;Note: the source is designated as rhesus monkey
C;Genetics:
A;Gene: PROS
C;Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat homo
F;1-51/Domain: Gla domain homology (fragment) <GLA>
F;1-7/Domain: signal sequence (fragment) #status predicted <SIG>
F;8-642/Product: plasma protein S #status predicted <MAT>
F;87-120/Domain: EGF homology <EG1>
F;127-165/Domain: EGF homology <EG2>
```

F;171-207/Domain: EGF homology <EG3>
F;213-248/Domain: EGF homology <EG4>
F;281-633/Domain: sex hormone-binding globulin homology <SHB>
F;291-444/Domain: laminin G repeat homology <LGR>

Query Match 43.1%; Score 85; DB 2; Length 642;
Best Local Similarity 38.6%; Pred. No. 1.5e-05;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||.|   :||:| | ||  :|:   |: :||:  :|  |: |:
Db          8 ANSMLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 51
```

RESULT 15
KXHUS
plasma protein S precursor - human
N;Alternate names: vitamin K-dependent protein S
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1990 #sequence_revision 26-Jan-1996 #text_change 16-Jul-1999
C;Accession: A35610; A35611; A26157; A25891; A35612; A60903; S02424; S09519
R;Schmidel, D.K.; Tatro, A.V.; Phelps, L.G.; Tomczak, J.A.; Long, G.L.
Biochemistry 29, 7845-7852, 1990
A;Title: Organization of the human protein S genes.
A;Reference number: A35610; MUID:91084444; PMID:2148110
A;Accession: A35610
A;Molecule type: DNA
A;Residues: 1-676 <SCH>
A;Cross-references: GB:M57853; NID:g190547; PIDN:AAA60357.1; PID:g190549; GB:J02917
A;Note: the authors translated the codon TTT for residue 26 as Leu
R;Ploos van Amstel, H.K.; Reitsma, P.H.; van der Logt, C.P.E.; Bertina, R.M.
Biochemistry 29, 7853-7861, 1990
A;Title: Intron-exon organization of the active human protein S gene PSalpha and its pse
A;Reference number: A35611; MUID:91084445; PMID:2148111
A;Accession: A35611
A;Molecule type: DNA
A;Residues: 1-25 <PL3>
A;Cross-references: GB:J02918
R;Hoskins, J.; Norman, D.K.; Beckmann, R.J.; Long, G.L.
Proc. Natl. Acad. Sci. U.S.A. 84, 349-353, 1987
A;Title: Cloning and characterization of human liver cDNA encoding a protein S precursor
A;Reference number: A26157; MUID:87092407; PMID:3467362
A;Accession: A26157
A;Molecule type: mRNA
A;Residues: 1-10,'P',12-25,'L',27-676 <HOS>
A;Cross-references: GB:M15036; NID:g190288; PIDN:AAA36479.1; PID:g190289
R;Lundwall, A.; Dackowski, W.; Cohen, E.; Shaffer, M.; Mahr, A.; Dahlback, B.; Stenflo,
Proc. Natl. Acad. Sci. U.S.A. 83, 6716-6720, 1986
A;Title: Isolation and sequence of the cDNA for human protein S, a regulator of blood co
A;Reference number: A25891; MUID:86313649; PMID:2944113
A;Accession: A25891
A;Molecule type: mRNA
A;Residues: 27-220,'L',222-262,'H',264-344,'Y',346-676 <LUN>
A;Cross-references: GB:M14338; NID:g190448; PIDN:AAA60181.1; PID:g190449
A;Note: part of this sequence, including the amino end of the mature protein, was determ
R;Edenbrandt, C.M.; Lundwall, A.; Wydro, R.; Stenflo, J.
Biochemistry 29, 7861-7868, 1990
A;Title: Molecular analysis of the gene for vitamin K dependent protein S and its pseudo
A;Reference number: A35612; MUID:91084446; PMID:2148112
A;Accession: A35612
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 284-676 <EDE>
A;Cross-references: GB:J02919
R;Ploos van Amstel, J.K.; van der Zanden, A.L.; Bakker, E.; Reitsma, P.H.; Bertina, R.M.
Thromb. Haemost. 58, 982-987, 1987
A;Title: Two genes homologous with human protein S cDNA are located on chromosome 3.
A;Reference number: A60903; MUID:88178564; PMID:2895503
A;Accession: A60903
A;Molecule type: mRNA
A;Residues: 351-676 <PLO>
R;Ploos van Amstel, H.K.; van der Zanden, A.L.; Reitsma, P.H.; Bertina, R.M.
FEBS Lett. 222, 186-190, 1987

A;Title: Human protein S cDNA encodes Phe-16 and Tyr 222 in consensus sequences for the
A;Reference number: S02424; MUID:88005138; PMID:2820795
A;Accession: S02424
A;Molecule type: mRNA
A;Residues: 1-676 <PL2>
A;Cross-references: EMBL:Y00692; NID:g36578; PIDN:CAA68687.1; PID:g36579
C;Genetics:
A;Gene: GDB:PROS1; PROS
A;Cross-references: GDB:120721; OMIM:176880
A;Map position: 3p11.1-3q11.2
A;Introns: 26/1; 78/3; 87/1; 116/1; 157/1; 201/1; 243/1; 283/3; 322/2; 385/3; 441/3; 498
C;Complex: in plasma forms a complex with C4b binding protein
C;Function:
A;Description: a cofactor for activated protein C (EC 3.4.21.69); thrombin cleavage dest
C;Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat hom
C;Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coagulation; carbox
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-41/Domain: propeptide #status predicted <PRO>
F;26-85/Domain: Gla domain homology <GLA>
F;42-676/Product: plasma protein S #status predicted <MAT>
F;121-154/Domain: EGF homology <EG1>
F;161-199/Domain: EGF homology <EG2>
F;205-241/Domain: EGF homology <EG3>
F;247-282/Domain: EGF homology <EG4>
F;315-667/Domain: sex hormone-binding globulin homology <SHB>
F;325-478/Domain: laminin G repeat homology <LGR>
F;47,48,55,57,60,61,66,67,70,73,77/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;58-63,88-113,121-134,126-143,145-154,161-175,171-184,186-199,205-217,212-226,228-241,2
F;111-112/Cleavage site: Arg-Ser (thrombin) #status predicted
F;136/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;177,219,258/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;499,509,530/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.1%; Score 85; DB 1; Length 676;
Best Local Similarity 38.6%; Pred. No. 1.6e-05;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||.|   :||:| | ||  :|:   |: :||:  :|  |: |:
Db         42 ANSLLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 85
```

Search completed: March 1, 2004, 10:10:01
Job time : 14.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 1, 2004, 09:58:27 ; Search time 10.5 Seconds
(without alignments)
218.199 Million cell updates/sec

Title: SEQ1-4SUBS
Perfect score: 197
Sequence: 1 ANSFLXXLRqgSLXRXCIXX.....XXAKXIFedVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	160	81.2	461	1	PRTC_HUMAN	P04070 homo sapien
2	140	71.1	461	1	PRTC_MOUSE	P33587 mus musculu
3	139	70.6	461	1	PRTC_RAT	P31394 rattus norv
4	138	70.1	458	1	PRTC_RABIT	Q28661 oryctolagus
5	123	62.4	459	1	PRTC_PIG	Q9glp2 sus scrofa
6	122	61.9	456	1	PRTC_BOVIN	P00745 bos taurus
7	114	57.9	492	1	FA10_BOVIN	P00743 bos taurus
8	110	55.8	488	1	FA10_HUMAN	P00742 homo sapien
9	107	54.3	231	1	TMG3_HUMAN	Q9bzd7 homo sapien
10	103	52.3	490	1	FA10_RABIT	O19045 oryctolagus
11	101	51.3	444	1	FA7_RABIT	P98139 oryctolagus
12	99	50.3	466	1	FA7_HUMAN	P08709 homo sapien
13	92	46.7	218	1	TMG1_HUMAN	O14668 homo sapien
14	86.5	43.9	617	1	THRB_RAT	P18292 rattus norv
15	86.5	43.9	618	1	THRB_MOUSE	P19221 mus musculu
16	86	43.7	475	1	FA10_CHICK	P25155 gallus gall
17	85	43.1	407	1	FA7_BOVIN	P22457 bos taurus
18	85	43.1	649	1	PRTS_MACMU	Q28520 macaca mula
19	85	43.1	676	1	PRTS_HUMAN	P07225 homo sapien
20	84.5	42.9	226	1	TMG4_HUMAN	Q9bzd6 homo sapien
21	84	42.6	376	1	FA10_HOPST	P83370 hoplocephal
22	84	42.6	622	1	THRB_HUMAN	P00734 homo sapien
23	82	41.6	376	1	FA10_TROCA	P81428 tropidechis
24	81	41.1	646	1	PRTS_RABIT	P98118 oryctolagus
25	80	40.6	446	1	FA7_MOUSE	P70375 mus musculu
26	80	40.6	452	1	FA9_CANFA	P19540 canis famil
27	80	40.6	459	1	FA9_MOUSE	P16294 mus musculu
28	80	40.6	461	1	FA9_HUMAN	P00740 homo sapien
29	80	40.6	461	1	FA9_PANTR	Q95nd7 pan troglod
30	80	40.6	675	1	PRTS_BOVIN	P07224 bos taurus
31	78	39.6	675	1	PRTS_RAT	P53813 rattus norv
32	73	37.1	416	1	FA9_BOVIN	P00741 bos taurus
33	72	36.5	98	1	FA10_NOTSC	P82807 notechis sc
34	72	36.5	625	1	THRB_BOVIN	P00735 bos taurus
35	71	36.0	675	1	PRTS_MOUSE	Q08761 mus musculu
36	69.5	35.3	396	1	PRTZ_BOVIN	P00744 bos taurus
37	65.5	33.2	400	1	PRTZ_HUMAN	P22891 homo sapien
38	65	33.0	202	1	TMG2_HUMAN	O14669 homo sapien
39	52	26.4	501	1	MKC1_CANAL	P43068 candida alb
40	50	25.4	363	1	ADK_TOXGO	Q9tvw2 toxoplasma
41	49.5	25.1	941	1	VP11_HUMAN	Q9h270 homo sapien
42	49.5	25.1	941	1	VP11_MOUSE	Q91w86 mus musculu
43	49	24.9	1363	1	VGR3_MOUSE	P35917 mus musculu
44	49	24.9	2812	1	ZAN_HUMAN	Q9y493 homo sapien
45	48	24.4	422	1	SPM1_SCHPO	Q92398 schizosacch

ALIGNMENTS

RESULT 1
PRTC_HUMAN
ID PRTC_HUMAN STANDARD; PRT; 461 AA.
AC P04070; Q15189; Q15190; Q16001;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE factor XIV).
GN PROC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270390; PubMed=2991887;
RA Foster D.C., Yoshitake S., Davie E.W.;
RT "The nucleotide sequence of the gene for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85269639; PubMed=2991859;
RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutzky J., Crabtree G.R.,
RA Long G.L.;
RT "The structure and evolution of a 461 amino acid human protein C
RT precursor and its messenger RNA, based upon the DNA sequence of
RT cloned human liver cDNAs.";
RL Nucleic Acids Res. 13:5233-5247(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86120978; PubMed=3511471;
RA Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;
RT "Evolution and organization of the human protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 106-461 FROM N.A.
RX MEDLINE=84272714; PubMed=6589623;
RA Foster D.C., Davie E.W.;
RT "Characterization of a cDNA coding for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-371.
RX MEDLINE=90293094; PubMed=1694179;
RA Miletich J.P., Broze G.J. Jr.;
RT "Beta protein C is not glycosylated at asparagine 329. The rate of
RT translation may influence the frequency of usage at asparagine-X-
RT cysteine sites.";
RL J. Biol. Chem. 265:11397-11404(1990).
RN [8]
RP HYDROXYLATION.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [9]
RP 3D-STRUCTURE MODELING OF 175-450.
RX MEDLINE=94272342; PubMed=8003977;
RA Fisher C.L., Greengard J.S., Griffin J.H.;
RT "Models of the serine protease domain of the human antithrombotic
RT plasma factor activated protein C and its zymogen.";
RL Protein Sci. 3:588-599(1994).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.
RX MEDLINE=97157472; PubMed=9003757;
RA Mather T., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,
RA Bode W.;
RT "The 2.8 A crystal structure of Gla-domainless activated protein C.";
RL EMBO J. 15:6822-6831(1996).
RN [11]
RP REVIEW ON PROC VARIANTS.
RX MEDLINE=93190290; PubMed=8446940;
RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,
RA Sala N., Cooper D.N.;
RT "Protein C deficiency: a database of mutations. For the Protein C & S
RT Subcommittee of the Scientific and Standardization Committee of the
RT International Society on Thrombosis and Haemostasis.";
RL Thromb. Haemost. 69:77-84(1993).
RN [12]
RP VARIANT PROC DEFICIENCY CYS-444.
RX MEDLINE=87204221; PubMed=2437584;
RA Romeo G., Hassan H.J., Staempfli S., Roncuzzi L., Cianetti L.,
RA Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,
RA Cortese R.;
RT "Hereditary thrombophilia: identification of nonsense and missense
RT mutations in the protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).
RN [13]
RP VARIANT PROC DEFICIENCY TRP-211.
RX MEDLINE=90098906; PubMed=2602169;
RA Grundy C.B., Chitolie A., Talbot S., Bevan D., Kakkar V.V.,
RA Cooper D.N.;
RT "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in
RT the protein C gene causing thrombosis.";
RL Nucleic Acids Res. 17:10513-10513(1989).
RN [14]
RP VARIANT PROC DEFICIENCY CYS-272.
RX MEDLINE=91329836; PubMed=1868249;
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;
RT "The spectrum of genetic defects in a panel of 40 Dutch families with
RT symptomatic protein C deficiency type I: heterogeneity and founder
RT effects.";
RL Blood 78:890-894(1991).
RN [15]
RP VARIANTS PROC DEFICIENCY ALA-62 AND MET-76.
RX MEDLINE=92190481; PubMed=1347706;
RA Bovill E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,
RA Rainville I.R., Long G.L.;
RT "Protein CVermont: symptomatic type II protein C deficiency
RT associated with two GLA domain mutations.";
RL Blood 79:1456-1465(1992).
RN [16]
RP VARIANT PROC DEFICIENCY ASP-418.
RX MEDLINE=92305321; PubMed=1611081;
RA Sugahara Y., Miura O., Yuen P., Aoki N.;
RT "Protein C deficiency Hong Kong 1 and 2: hereditary protein C
RT deficiency caused by two mutant alleles, a 5-nucleotide deletion and
RT a missense mutation.";
RL Blood 80:126-133(1992).
RN [17]
RP VARIANT PROC DEFICIENCY LEU-289.
RX MEDLINE=92380660; PubMed=1511988;
RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;
RT "A novel homozygous missense mutation in the protein C (PROC) gene
RT causing recurrent venous thrombosis.";
RL Hum. Genet. 89:683-684(1992).
RN [18]
RP VARIANTS PROC DEFICIENCY GLN-220 AND TRP-220.
RX MEDLINE=92380661; PubMed=1511989;
RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;
RT "Two different missense mutations at Arg 178 of the protein C (PROC)
RT gene causing recurrent venous thrombosis.";
RL Hum. Genet. 89:685-686(1992).
RN [19]
RP VARIANT PROC DEFICIENCY GLN-220.
RX MEDLINE=93250852; PubMed=1301959;
RA Gandrille S., Vidaud M., Aiach M., Alhenc-Gelas M., Fischer A.M.,
RA Gouault-Heilman M., Toulon P., Fiessinger J.N., Goossens M.;
RT "Two novel mutations responsible for hereditary type I protein C
RT deficiency: characterization by denaturing gradient gel
RT electrophoresis.";
RL Hum. Mutat. 1:491-500(1992).
RN [20]
RP VARIANT PROC DEFICIENCY SER-334.
RX MEDLINE=92276939; PubMed=1593215;
RA Yamamoto K., Matsushita T., Sugiura I., Takamatsu J., Iwasaki E.,
RA Wada H., Deguchi K., Shirakawa S., Saito H.;
RT "Homozygous protein C deficiency: identification of a novel missense
RT mutation that causes impaired secretion of the mutant protein C.";
RL J. Lab. Clin. Med. 119:682-689(1992).
RN [21]
RP VARIANTS PROC DEFICIENCY TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.
RX MEDLINE=93313192; PubMed=8324221;
RA Gandrille S., Alhenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,
RA Juhan-Vague I., Aiach M.;
RT "Five novel mutations located in exons III and IX of the protein C
RT gene in patients presenting with defective protein C anticoagulant
RT activity.";
RL Blood 82:159-168(1993).
RN [22]
RP VARIANTS PROC DEFICIENCY GLY-14; GLN-211; TYR-244; GLN-253; LEU-321;
RP CYS-328; ILE-385; THR-388 AND VAL-388.
RX MEDLINE=93271391; PubMed=8499565;
RA Poort S.R., Pabinger-Fasching I., Mannhalter C., Reitsma P.H.,
RA Bertina R.M.;
RT "Twelve novel and two recurrent mutations in 14 Austrian families
RT with hereditary protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:273-280(1993).
RN [23]
RP VARIANT PROC DEFICIENCY TRP-57.
RX MEDLINE=93271396; PubMed=8499568;

```
  Query Match           81.2%;  Score 160;  DB 1;  Length 461;
  Best Local Similarity  70.5%;  Pred. No. 2.8e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  || | ||  ||||  || ||::|||||||||||
Db         43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
PRTC_MOUSE
ID   PRTC_MOUSE     STANDARD;      PRT;   461 AA.
AC   P33587; O35498;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92316897; PubMed=1618739;
RA   Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;
RT   "Isolation and characterization of a mouse protein C cDNA.";
RL   J. Biochem. 111:491-495(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98152576; PubMed=9493582;
RA   Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,
RA   Castellino F.J.;
RT   "Nucleotide structure and characterization of the murine gene encoding
RT   anticoagulant protein C.";
RL   Thromb. Haemost. 79:310-316(1998).
RN   [3]
RP   SEQUENCE OF 274-434 FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=94318474; PubMed=8043441;
RA   Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT   "A comparative study of partial primary structures of the catalytic
RT   region of mammalian protein C.";
RL   Br. J. Haematol. 86:590-600(1994).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10445; BAA01235.1; -.
DR   EMBL; AF034569; AAC33795.1; -.
DR   EMBL; D43755; BAA07812.1; -.
DR   PIR; JX0210; JX0210.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   MGD; MGI:97771; Proc.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     33       BY SIMILARITY.
FT   PROPEP      34     41       BY SIMILARITY.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN      199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE    199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     213    461       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   253    253       CHARGE RELAY SYSTEM.
FT   ACT_SITE   299    299       CHARGE RELAY SYSTEM.
FT   ACT_SITE   402    402       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
```

```
FT   DISULFID    100    105       BY SIMILARITY.
FT   DISULFID    104    119       BY SIMILARITY.
FT   DISULFID    121    130       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
FT   DISULFID    146    159       BY SIMILARITY.
FT   DISULFID    161    174       BY SIMILARITY.
FT   DISULFID    182    319       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    238    254       BY SIMILARITY.
FT   DISULFID    373    387       BY SIMILARITY.
FT   DISULFID    398    426       BY SIMILARITY.
FT   CARBOHYD    214    214       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    328    328       MISSING (IN REF. 2).
FT   CONFLICT    393    393       N -> D (IN REF. 2).
SQ   SEQUENCE   461 AA;  51945 MW;  53FAA0D85B194D6E CRC64;

  Query Match             71.1%;  Score 140;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 7e-16;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db         42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
PRTC_RAT
ID   PRTC_RAT       STANDARD;      PRT;   461 AA.
AC   P31394;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=92329550; PubMed=1627650;
RA   Okafuji T., Maekawa K., Nawa K., Marumoto Y.;
RT   "The cDNA cloning and mRNA expression of rat protein C.";
RL   Biochim. Biophys. Acta 1131:329-332(1992).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X64336; CAA45617.1; -.
DR   PIR; S18994; S18994.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL        1     32       BY SIMILARITY.
FT   PROPEP       33     41       BY SIMILARITY.
FT   CHAIN        42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN       199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE     199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE        212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN       96    131       EGF-LIKE 1.
FT   DOMAIN      135    175       EGF-LIKE 2.
FT   DOMAIN      213    461       SERINE PROTEASE.
FT   MOD_RES      47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES     112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE    254    254       CHARGE RELAY SYSTEM.
FT   ACT_SITE    300    300       CHARGE RELAY SYSTEM.
FT   ACT_SITE    402    402       CHARGE RELAY SYSTEM.
FT   DISULFID     58     63       BY SIMILARITY.
FT   DISULFID     91    110       BY SIMILARITY.
FT   DISULFID    100    105       BY SIMILARITY.
FT   DISULFID    104    119       BY SIMILARITY.
FT   DISULFID    121    130       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
```

```
FT   DISULFID    146    159       BY SIMILARITY.
FT   DISULFID    161    174       BY SIMILARITY.
FT   DISULFID    182    320       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    239    255       BY SIMILARITY.
FT   DISULFID    373    387       BY SIMILARITY.
FT   DISULFID    398    426       BY SIMILARITY.
FT   CARBOHYD    215    215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   461 AA;  51912 MW;  8A4CF93664EDACD5 CRC64;

  Query Match             70.6%;  Score 139;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 1e-15;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db         42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
PRTC_RABIT
ID   PRTC_RABIT     STANDARD;      PRT;   458 AA.
AC   Q28661;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Shen L., He X., Dahlback B.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49933; AAA92956.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL        <1     27       BY SIMILARITY.
FT   PROPEP        28     36       BY SIMILARITY.
FT   CHAIN         37     458      VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN         37     192      PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN        195     458      PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE      195     209      ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE         209     210      CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN        91     126      EGF-LIKE 1.
FT   DOMAIN       130     170      EGF-LIKE 2.
FT   DOMAIN       210     458      SERINE PROTEASE.
FT   MOD_RES       42      42      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       43      43      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       50      50      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       52      52      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       55      55      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       56      56      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       61      61      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       62      62      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES       65      65      GAMMA-CARBOXYGLUTAMIC ACID
FT                                 (BY SIMILARITY).
FT   MOD_RES      107     107      HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE     250     250      CHARGE RELAY SYSTEM.
FT   ACT_SITE     296     296      CHARGE RELAY SYSTEM.
FT   ACT_SITE     399     399      CHARGE RELAY SYSTEM.
FT   DISULFID      53      58      BY SIMILARITY.
FT   DISULFID      86     105      BY SIMILARITY.
FT   DISULFID      95     100      BY SIMILARITY.
FT   DISULFID      99     114      BY SIMILARITY.
FT   DISULFID     116     125      BY SIMILARITY.
FT   DISULFID     134     145      BY SIMILARITY.
FT   DISULFID     141     154      BY SIMILARITY.
FT   DISULFID     156     169      BY SIMILARITY.
FT   DISULFID     177     316      INTERCHAIN (BY SIMILARITY).
FT   DISULFID     235     251      BY SIMILARITY.
FT   DISULFID     370     384      BY SIMILARITY.
FT   DISULFID     395     423      BY SIMILARITY.
```

```
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    287    287       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   458 AA;  51087 MW;  D75A5F990C8F29D7 CRC64;

  Query Match             70.1%;  Score 138;  DB 1;  Length 458;
  Best Local Similarity   59.1%;  Pred. No. 1.5e-15;
  Matches   26;  Conservative    4;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  ||  | |:  :||   || ||: |||||||| |:
Db         37 ANSFLEELRPSSLERECVEEVCDLEEAKEIFQSVDDTLAFWYKY 80


RESULT 5
PRTC_PIG
ID   PRTC_PIG       STANDARD;      PRT;   459 AA.
AC   Q9GLP2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21121490; PubMed=11229814;
RA   Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
RA   Kim H.K.W.;
RT   "Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT   protein modeling of membrane binding sites and comparative anatomy of
RT   domains.";
RL   Cell. Mol. Life Sci. 58:148-159(2001).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF191307; AAG28380.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     18       BY SIMILARITY.
FT   PROPEP      19     41       BY SIMILARITY.
FT   CHAIN       42    459       VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY
FT                               SIMILARITY).
FT   CHAIN      199    459       PROTEIN C HEAVY CHAIN (BY
FT                               SIMILARITY).
FT   PEPTIDE    199    213       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       213    214       CLEAVAGE (BY THROMBIN) (BY
FT                               SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     214    459       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   255    255       CHARGE RELAY SYSTEM.
FT   ACT_SITE   301    301       CHARGE RELAY SYSTEM.
FT   ACT_SITE   400    400       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
FT   DISULFID   100    105       BY SIMILARITY.
FT   DISULFID   104    119       BY SIMILARITY.
FT   DISULFID   121    130       BY SIMILARITY.
FT   DISULFID   139    150       BY SIMILARITY.
FT   DISULFID   146    159       BY SIMILARITY.
FT   DISULFID   161    174       BY SIMILARITY.
FT   DISULFID   182    321       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   240    256       BY SIMILARITY.
FT   DISULFID   371    385       BY SIMILARITY.
```

```
FT   DISULFID    396    424       BY SIMILARITY.
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   459 AA;  51866 MW;  8541AAC14CC16D09 CRC64;

  Query Match             62.4%;  Score 123;  DB 1;  Length 459;
  Best Local Similarity   52.3%;  Pred. No. 5.4e-13;
  Matches   23;  Conservative    7;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ||  || | |     ||| |: ||:: ::|:|||||:
Db         42 ANSFLEELRPSSLERECKEETCDFEEAREIFQNTENTMAFWSKY 85


RESULT 6
PRTC_BOVIN
ID   PRTC_BOVIN     STANDARD;      PRT;   456 AA.
AC   P00745;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vitamin-K-dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85014826; PubMed=6091100;
RA   Long G.L., Balagaje R.M., McGillivray R.T.A.;
RT   "Cloning and sequencing of liver cDNA coding for bovine protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).
RN   [2]
RP   SEQUENCE OF 40-194, AND CARBOHYDRATE-LINKAGE SITE ASN-136.
RX   MEDLINE=83007325; PubMed=6896876;
RA   Fernlund P., Stenflo J.;
RT   "Amino acid sequence of the light chain of bovine protein C.";
RL   J. Biol. Chem. 257:12170-12179(1982).
RN   [3]
RP   REVISION TO 110.
RX   MEDLINE=83169769; PubMed=6572939;
RA   Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;
RT   "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).
RN   [4]
RP   SEQUENCE OF 197-456, AND CARBOHYDRATE-LINKAGE SITES ASN-289; ASN-350
RP   AND ASN-366.
RX   MEDLINE=83007326; PubMed=6896877;
RA   Stenflo J., Fernlund P.;
RT   "Amino acid sequence of the heavy chain of bovine protein C.";
RL   J. Biol. Chem. 257:12180-12190(1982).
RN   [5]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213513; PubMed=6304092;
RA   Esmon N.L., Debault L.E., Esmon C.T.;
RT   "Proteolytic formation and properties of gamma-carboxyglutamic acid-
RT   domainless protein C.";
RL   J. Biol. Chem. 258:5548-5553(1983).
RN   [6]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213514; PubMed=6406503;
RA   Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;
RT   "Structural changes required for activation of protein C are induced
RT   by Ca2+ binding to a high affinity site that does not contain gamma-
RT   carboxyglutamic acid.";
RL   J. Biol. Chem. 258:5554-5560(1983).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the thrombin-
CC       thrombomodulin complex.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02435; AAA30685.1; -.
DR   PIR; A26250; KXBO.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     29
FT   PROPEP        30     39
FT   CHAIN         40    194       PROTEIN C LIGHT CHAIN.
FT   CHAIN        197    456       PROTEIN C HEAVY CHAIN.
FT   PEPTIDE      197    210       ACTIVATION PEPTIDE.
FT   DOMAIN        94    129       EGF-LIKE 1.
FT   DOMAIN       133    173       EGF-LIKE 2.
FT   DOMAIN       211    456       SERINE PROTEASE.
FT   MOD_RES       45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       55     55       GAMMA-CARBOXYGLUTAMIC ACID.
```

```
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      62     62       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     110    110       HYDROXYLATION.
FT   ACT_SITE    252    252       CHARGE RELAY SYSTEM.
FT   ACT_SITE    298    298       CHARGE RELAY SYSTEM.
FT   ACT_SITE    397    397       CHARGE RELAY SYSTEM.
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    108       BY SIMILARITY.
FT   DISULFID     98    103       BY SIMILARITY.
FT   DISULFID    102    117       BY SIMILARITY.
FT   DISULFID    119    128       BY SIMILARITY.
FT   DISULFID    137    148       BY SIMILARITY.
FT   DISULFID    144    157       BY SIMILARITY.
FT   DISULFID    159    172       BY SIMILARITY.
FT   DISULFID    180    318       INTERCHAIN.
FT   DISULFID    237    253
FT   DISULFID    368    382
FT   DISULFID    393    421
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .).
FT   VARIANT      82     82       F -> K.
FT   CONFLICT    455    456       VP -> PV (IN REF. 4).
SQ   SEQUENCE   456 AA;  51407 MW;  CAAF6833F894C209 CRC64;

  Query Match             61.9%;  Score 122;  DB 1;  Length 456;
  Best Local Similarity   50.0%;  Pred. No. 7.9e-13;
  Matches   21;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
              |||||  || |:: | |   :|:|  |: ||:: :||:|||| 
Db         40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 7
FA10_BOVIN
ID   FA10_BOVIN     STANDARD;      PRT;   492 AA.
AC   P00743;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE OF 1-487 FROM N.A.
RX   MEDLINE=84247315; PubMed=6330671;
RA   Fung M.R., Campbell R.M., McGillivray R.T.A.;
RT   "Blood coagulation factor X mRNA encodes a single polypeptide chain
RT   containing a prepro leader sequence.";
RL   Nucleic Acids Res. 12:4481-4492(1984).
RN   [2]
RP   SEQUENCE OF 41-180.
RX   MEDLINE=80130563; PubMed=6766735;
RA   Enfield D.L., Ericsson L.H., Fujikawa K., Walsh K.A., Neurath H.,
RA   Titani K.;
RT   "Amino acid sequence of the light chain of bovine factor X1 (Stuart
RT   factor).";
RL   Biochemistry 19:659-667(1980).
RN   [3]
RP   REVISION TO 103.
RX   MEDLINE=83308813; PubMed=6688526;
RA   McMullen B.A., Fujikawa K., Kisiel W.;
RT   "The occurrence of beta-hydroxyaspartic acid in the vitamin
RT   K-dependent blood coagulation zymogens.";
RL   Biochem. Biophys. Res. Commun. 115:8-14(1983).
RN   [4]
RP   SEQUENCE OF 183-492, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX   MEDLINE=76053069; PubMed=1059093;
RA   Titani K., Fujikawa K., Enfield D.L., Ericsson L.H., Walsh K.A.,
RA   Neurath H.;
RT   "Bovine factor X1 (Stuart factor): amino-acid sequence of heavey
RT   chain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3082-3086(1975).
RN   [5]
RP   SEQUENCE OF 183-233, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=94062825; PubMed=8243461;
RA   Inoue K., Morita T.;
RT   "Identification of O-linked oligosaccharide chains in the activation
RT   peptides of blood coagulation factor X. The role of the carbohydrate
RT   moieties in the activation of factor X.";
RL   Eur. J. Biochem. 218:153-163(1993).
RN   [6]
RP   ACTIVE SITE.
RX   MEDLINE=73053314; PubMed=4264286;
RA   Titani K., Hermodson M.A., Fujikawa K., Ericsson L.H., Walsh K.A.,
RA   Neurath H., Davie E.W.;
RT   "Bovine factor X 1a (activated Stuart factor). Evidence of homology
RT   with mammalian serine proteases.";
RL   Biochemistry 11:4899-4903(1972).
RN   [7]
RP   PROCESSING.
RX   MEDLINE=76053121; PubMed=1059122;
RA   Fujikawa K., Titani K., Davie E.W.;
RT   "Activation of bovine factor X (Stuart factor): conversion of factor
RT   Xa-alpha to factor Xa-beta.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3359-3363(1975).
RN   [8]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=84185716; PubMed=6546930;
RA   Sugo T., Bjoerk I., Holmgren A., Stenflo J.;
RT   "Calcium-binding properties of bovine factor X lacking the gamma-
RT   carboxyglutamic acid-containing region.";
RL   J. Biol. Chem. 259:5705-5710(1984).
RN   [9]
RP   SULFATION.
RX   MEDLINE=86140210; PubMed=3949800;
RA   Morita T., Jackson C.M.;
RT   "Localization of the structural difference between bovine blood
RT   coagulation factors X1 and X2 to tyrosine 18 in the activation
RT   peptide.";
RL   J. Biol. Chem. 261:4008-4014(1986).
RN   [10]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=91084483; PubMed=2261466;
RA   Selander M., Persson E., Stenflo J., Drakenberg T.;
RT   "1H NMR assignment and secondary structure of the Ca2(+)-free form of
RT   the amino-terminal epidermal growth factor like domain in coagulation
RT   factor X.";
RL   Biochemistry 29:8111-8118(1990).
RN   [11]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92329412; PubMed=1627540;
RA   Ullner M., Selander M., Persson E., Stenflo J., Drakenberg T.,
RA   Teleman O.;
RT   "Three-dimensional structure of the apo form of the N-terminal
RT   EGF-like module of blood coagulation factor X as determined by NMR
RT   spectroscopy and simulated folding.";
RL   Biochemistry 31:5974-5983(1992).
RN   [12]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92406922; PubMed=1527084;
RA   Selander-Sunnerhagen M., Ullner M., Persson E., Teleman O.,
RA   Stenflo J., Drakenberg T.;
RT   "How an epidermal growth factor (EGF)-like domain binds calcium. High
RT   resolution NMR structure of the calcium form of the NH2-terminal EGF-
```

```
RT   like domain in coagulation factor X.";
RL   J. Biol. Chem. 267:19642-19649(1992).
RN   [13]
RP   STRUCTURE BY NMR OF 41-126.
RX   MEDLINE=96387194; PubMed=8794734;
RA   Sunnerhagen M., Olah G.A., Stenflo J., Forsen S., Drakenberg T.,
RA   Trewhella J.;
RT   "The relative orientation of Gla and EGF domains in coagulation
RT   factor X is altered by Ca2+ binding to the first EGF domain. A
RT   combined NMR-small angle X-ray scattering study.";
RL   Biochemistry 35:11547-11559(1996).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC       by the excision of two Arg residues and are held together by 1 or
CC       more disulfide bonds.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium.
CC   -!- PTM: N- and O-glycosylated.
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X00673; CAA25286.1; -.
DR   PIR; A22867; EXBO.
DR   PDB; 1APO; 31-JAN-94.
DR   PDB; 1CCF; 31-MAY-94.
DR   PDB; 1WHE; 15-MAY-97.
DR   PDB; 1WHF; 15-MAY-97.
DR   PDB; 1IOD; 21-JAN-03.
DR   PDB; 1KIG; 28-OCT-98.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00743; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat; Sulfation; 3D-structure.
FT   SIGNAL       1     23       POTENTIAL.
FT   PROPEP      24     40
FT   CHAIN       41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN      183    492       FACTOR X HEAVY CHAIN.
FT   PROPEP     183    233       ACTIVATION PEPTIDE.
FT   CHAIN      234    492       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   PROPEP     476    492       MAY BE REMOVED BUT IS NOT NECESSARY FOR
FT                               ACTIVATION.
FT   DOMAIN      86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125    165       EGF-LIKE 2.
FT   DOMAIN     234    492       SERINE PROTEASE.
FT   ACT_SITE   275    275       CHARGE RELAY SYSTEM.
FT   ACT_SITE   321    321       CHARGE RELAY SYSTEM.
FT   ACT_SITE   418    418       CHARGE RELAY SYSTEM.
FT   MOD_RES     46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     54     54       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     56     56       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID.

  Query Match             57.9%;  Score 114;  DB 1;  Length 492;
  Best Local Similarity   45.5%;  Pred. No. 2e-11;
  Matches   20;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  ::||:| | |:   |    |: :||| : |  ||||:
Db         41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84


RESULT 8
FA10_HUMAN
ID   FA10_HUMAN     STANDARD;      PRT;   488 AA.
AC   P00742; Q14340;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91216473; PubMed=1902434;
RA   Messier T.L., Pittman D.D., Long G.L., Kaufman R.J., Church W.R.;
RT   "Cloning and expression in COS-1 cells of a full-length cDNA encoding
RT   human coagulation factor X.";
RL   Gene 99:291-294(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87026600; PubMed=3768336;
RA   Leytus S.P., Foster D.C., Kurachi K., Davie E.W.;
RT   "Gene for human factor X: a blood coagulation factor whose gene
RT   organization is essentially identical with that of factor IX and
RT   protein C.";
RL   Biochemistry 25:5098-5102(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 13-488 FROM N.A.
RX   MEDLINE=85216545; PubMed=2582420;
RA   Fung M.R., Hay C.W., McGillivray R.T.A.;
RT   "Characterization of an almost full-length cDNA coding for human
RT   blood coagulation factor X.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3591-3595(1985).
RN   [6]
RP   SEQUENCE OF 19-488 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=86221713; PubMed=3011603;
RA   Kaul R.K., Hildebrand B., Roberts S., Jagadeeswaran P.;
RT   "Isolation and characterization of human blood-coagulation factor X
RT   cDNA.";
RL   Gene 41:311-314(1986).
RN   [7]
RP   SEQUENCE OF 41-179.
RX   MEDLINE=83257207; PubMed=6871167;
RA   McMullen B.A., Fujikawa K., Kisiel W., Sasagawa T., Howald W.N.,
RA   Kwa E.Y., Weinstein B.;
RT   "Complete amino acid sequence of the light chain of human blood
RT   coagulation factor X: evidence for identification of residue 63 as
RT   beta-hydroxyaspartic acid.";
RL   Biochemistry 22:2875-2884(1983).
RN   [8]
RP   SEQUENCE OF 115-488 FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Liver;
RX   MEDLINE=84222026; PubMed=6587384;
RA   Leytus S.P., Chung D.W., Kisiel W., Kurachi K., Davie E.W.;
RT   "Characterization of a cDNA coding for human factor X.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:3699-3702(1984).
RN   [9]
RP   SEQUENCE OF 183-234, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=94062825; PubMed=8243461;
RA   Inoue K., Morita T.;
RT   "Identification of O-linked oligosaccharide chains in the activation
RT   peptides of blood coagulation factor X. The role of the carbohydrate
RT   moieties in the activation of factor X.";
RL   Eur. J. Biochem. 218:153-163(1993).
RN   [10]
RP   SEQUENCE OF 1-23 FROM N.A.
RX   MEDLINE=90128299; PubMed=2612918;
RA   Jagadeeswaran P., Reddy S.V., Rao K.J., Hamsabhushanam K., Lyman G.;
RT   "Cloning and characterization of the 5' end (exon 1) of the gene
RT   encoding human factor X.";
RL   Gene 84:517-519(1989).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-179 AND 235-278.
RX   MEDLINE=93360277; PubMed=8355279;
RA   Padmanabhan K., Padmanabhan K.P., Tulinsky A., Park C.H., Bode W.,
RA   Huber R., Blankenship D.T., Cardin A.D., Kisiel W.;
RT   "Structure of human des(1-45) factor Xa at 2.2-A resolution.";
RL   J. Mol. Biol. 232:947-966(1993).
RN   [12]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-179 AND 235-278.
RX   MEDLINE=98283982; PubMed=9618463;
RA   Kamata K., Kawamoto H., Honma T., Iwama T., Kim S.H.;
RT   "Structural basis for chemical inhibition of human blood coagulation
RT   factor Xa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6630-6635(1998).
RN   [13]
RP   VARIANTS ILE-7 AND HIS-30.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [14]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC       by the excision of two Arg residues and are held together by 1 or
CC       more disulfide bonds.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium.
CC   -!- PTM: N- and O-glycosylated.
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K03194; AAA52490.1; -.
DR   EMBL; M57285; AAA52421.1; -.
DR   EMBL; AF503510; AAM19347.1; -.
DR   EMBL; BC046125; AAH46125.1; -.
DR   EMBL; L29433; AAA52764.1; -.
DR   EMBL; L00390; AAA52764.1; JOINED.
DR   EMBL; L00391; AAA52764.1; JOINED.
DR   EMBL; L00392; AAA52764.1; JOINED.
DR   EMBL; L00393; AAA52764.1; JOINED.
DR   EMBL; L00394; AAA52764.1; JOINED.
DR   EMBL; L00395; AAA52764.1; JOINED.
DR   EMBL; L00396; AAA52764.1; JOINED.
DR   EMBL; M22613; AAA51984.1; -.
DR   EMBL; K01886; AAA52486.1; -.
DR   EMBL; M33297; AAA52636.1; -.
DR   PIR; A24478; EXHU.
DR   PDB; 1HCG; 08-MAY-95.
DR   PDB; 1FAX; 29-OCT-97.
DR   PDB; 1FXY; 17-JUN-98.
DR   PDB; 1XKA; 23-MAR-99.
DR   PDB; 1XKB; 23-MAR-99.
DR   PDB; 1EZQ; 20-SEP-00.
DR   PDB; 1F0R; 20-SEP-00.
DR   PDB; 1F0S; 20-SEP-00.
DR   PDB; 1FJS; 17-NOV-00.
DR   PDB; 1G2L; 20-OCT-01.
DR   PDB; 1G2M; 20-OCT-01.
```

```
DR   PDB; 1KSN; 19-JUN-02.
DR   PDB; 1KYE; 11-FEB-03.
DR   PDB; 1MQ5; 28-JAN-03.
DR   PDB; 1MQ6; 28-JAN-03.
DR   PDB; 1NFU; 25-FEB-03.
DR   PDB; 1NFW; 25-FEB-03.
DR   PDB; 1NFX; 25-FEB-03.
DR   PDB; 1NFY; 25-FEB-03.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00742; -.
DR   Genew; HGNC:3528; F10.
DR   MIM; 134530; -.
DR   MIM; 227600; -.
DR   GO; GO:0005576; C:extracellular; TAS.
DR   GO; GO:0003804; F:blood coagulation factor X activity; TAS.
DR   GO; GO:0007596; P:blood coagulation; TAS.

  Query Match             55.8%;  Score 110;  DB 1;  Length 488;
  Best Local Similarity   43.2%;  Pred. No. 9.3e-11;
  Matches   19;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  :::| | | |:   | :  |: :||| | |  ||:|:
Db         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84


RESULT 9
TMG3_HUMAN
ID   TMG3_HUMAN     STANDARD;      PRT;   231 AA.
AC   Q9BZD7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 3 precursor.
GN   TMG3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spinal cord;
RX   MEDLINE=21117044; PubMed=11171957;
RA   Kulman J.D., Harris J.E., Xie L., Davie E.W.;
RT   "Identification of two novel transmembrane gamma-carboxyglutamic acid
RT   proteins expressed broadly in fetal and adult tissues.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in brain, lung, kidney and heart.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamate by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF326350; AAK00955.1; -.
DR   HSSP; P00740; 1CFH.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     19       POTENTIAL.
FT   CHAIN          20    231       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 3.
FT   DOMAIN         20     78       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       79    101       POTENTIAL.
FT   DOMAIN        102    231       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         23     60       GLA-RICH.
SQ   SEQUENCE   231 AA;  25848 MW;  8A373E4848490D81 CRC64;

  Query Match             54.3%;  Score 107;  DB 1;  Length 231;
  Best Local Similarity   43.9%;  Pred. No. 1.3e-10;
  Matches   18;  Conservative    8;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              || ||  ||||:: | |:  || :   | :||: : |: ||
Db         20 ANEFLEELRQGTIERECMEEICSYEEVKEVFENKEKTMEFW 60


RESULT 10
FA10_RABIT
ID   FA10_RABIT     STANDARD;      PRT;   490 AA.
AC   O19045;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97256311; PubMed=9101642;
RA   Pendurthi U.R., Anderson K.D., James H.L.;
RT   "Characterization of a full-length cDNA for rabbit factor X.";
RL   Thromb. Res. 85:503-514(1997).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: The two chains are formed from a single-chain precursor
CC       by the excision of two Arg residues and are held together by 1 or
CC       more disulfide bonds.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium (By
CC       similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC       (BY SIMILARITY).
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003200; AAB62542.1; -.
DR   HSSP; P00742; 1HCG.
DR   MEROPS; S01.216; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
```

```
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat.
FT   SIGNAL       1     20       POTENTIAL.
FT   PROPEP      21     40       BY SIMILARITY.
FT   CHAIN       41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN      184    490       FACTOR X HEAVY CHAIN.
FT   PROPEP     184    232       ACTIVATION PEPTIDE.
FT   CHAIN      233    490       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   DOMAIN      86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125    165       EGF-LIKE 2.
FT   DOMAIN     233    490       SERINE PROTEASE.
FT   MOD_RES     46     46       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     54     54       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     56     56       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     59     59       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     65     65       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     69     69       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     72     72       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     75     75       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     79     79       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES    103    103       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   274    274       CHARGE RELAY SYSTEM.
FT   ACT_SITE   320    320       CHARGE RELAY SYSTEM.
FT   ACT_SITE   417    417       CHARGE RELAY SYSTEM.
FT   DISULFID    90    101       BY SIMILARITY.
FT   DISULFID    95    110       BY SIMILARITY.
FT   DISULFID   112    121       BY SIMILARITY.
FT   DISULFID   129    140       BY SIMILARITY.
FT   DISULFID   136    149       BY SIMILARITY.
FT   DISULFID   151    164       BY SIMILARITY.
FT   DISULFID   172    340       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   239    244       BY SIMILARITY.
FT   DISULFID   259    275       BY SIMILARITY.
FT   DISULFID   388    402       BY SIMILARITY.
FT   DISULFID   413    441       BY SIMILARITY.
FT   CARBOHYD    61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   490 AA;  53965 MW;  3A39FA85AF2A6D11 CRC64;

  Query Match             52.3%;  Score 103;  DB 1;  Length 490;
  Best Local Similarity   43.2%;  Pred. No. 1.4e-09;
  Matches   19;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              |||||  |::|:| | |:   | :  |  :||| : |  ||:|:
Db         41 ANSFLEELKKGNLERECMEENCSYEEALEVFEDREKTNEFWNKY 84


RESULT 11
FA7_RABIT
ID   FA7_RABIT      STANDARD;      PRT;   444 AA.
AC   P98139; P79224;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator).
GN   F7.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93190306; PubMed=8383365;
RA   Brothers A.B., Clarke B.J., Sheffield W.P., Blajchman M.A.;
RT   "Complete nucleotide sequence of the cDNA encoding rabbit coagulation
RT   factor VII.";
RL   Thromb. Res. Suppl. 69:231-238(1993).
RN   [2]
RP   REVISION TO 395.
RC   TISSUE=Liver;
RA   Ruiz S.R., Blajchman M.A., Clarke B.J.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Circulates in the blood in a zymogen form. Factor VII is
CC       converted to factor VIIa by factor Xa, factor XIIa, factor IXa, or
CC       thrombin by minor proteolysis. In the presence of tissue factor
CC       and calcium ions, factor VIIa then converts factor X to factor Xa
CC       by limited proteolysis. Factor VIIa will also convert factor IX to
CC       factor IXa in the presence of tissue factor and calcium (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC       form factor Xa.
CC   -!- SUBUNIT: Heterodimer of a light chain and a heavy chain linked by
CC       a disulfide bond (By similarity).
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC       glutamate residues allows the modified protein to bind calcium (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U77477; AAB37326.1; -.
DR   HSSP; P08709; 1FAK.
DR   MEROPS; S01.215; -.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
```

```
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Blood coagulation; Zymogen; Glycoprotein;
KW   Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;
KW   EGF-like domain; Repeat; Signal; Hydroxylation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     39       POTENTIAL.
FT   CHAIN        40    191       FACTOR VII LIGHT CHAIN.
FT   CHAIN       192    444       FACTOR VII HEAVY CHAIN.
FT   DOMAIN       45     74       GLA-RICH.
FT   DOMAIN       85    121       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      126    167       EGF-LIKE 2.
FT   DOMAIN      192    444       SERINE PROTEASE.
FT   SITE        191    192       CLEAVAGE (BY FACTOR XA, FACTOR XIIA,
FT                                FACTOR IXA, OR THROMBIN) (BY SIMILARITY).
FT   ACT_SITE    232    232       BY SIMILARITY.
FT   ACT_SITE    281    281       BY SIMILARITY.
FT   ACT_SITE    383    383       BY SIMILARITY.
FT   BINDING     377    377       SUBSTRATE (BY SIMILARITY).
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    100       BY SIMILARITY.
FT   DISULFID     94    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    130    141       BY SIMILARITY.
FT   DISULFID    137    151       BY SIMILARITY.
FT   DISULFID    153    166       BY SIMILARITY.
FT   DISULFID    174    301       BY SIMILARITY.
FT   DISULFID    198    203       BY SIMILARITY.
FT   DISULFID    217    233       BY SIMILARITY.
FT   DISULFID    349    368       BY SIMILARITY.
FT   DISULFID    379    407       BY SIMILARITY.
FT   MOD_RES      45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     102    102       HYDROXYLATION (BY SIMILARITY).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   444 AA;  49011 MW;  0481ABC4FE5427F8 CRC64;

 Query Match             51.3%;  Score 101;  DB 1;  Length 444;
 Best Local Similarity   46.3%;  Pred. No. 2.8e-09;
 Matches   19;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              |||||  ||  |||  | |   :| |  |: :|:  : |  ||
Db         40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80


RESULT 12
FA7_HUMAN
ID   FA7_HUMAN      STANDARD;      PRT;   466 AA.
AC   P08709; Q14339;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator) (Eptacog alfa).
GN   F7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=86205965; PubMed=3486420;
RA   Hagen F.S., Gray C.L., O'Hara P.J., Grant F.J., Saari G.C.,
RA   Woodbury R.G., Hart C.E., Insley M.Y., Kisiel W., Kurachi K.,
RA   Davie E.W.;
RT   "Characterization of a cDNA coding for human factor VII.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:2412-2416(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87260948; PubMed=3037537;
RA   O'Hara P.J., Grant F.J., Haldeman B.A., Gray C.L., Insley M.Y.,
RA   Hagen F.S., Murray M.J.;
RT   "Nucleotide sequence of the gene coding for human factor VII, a
RT   vitamin K-dependent protein participating in blood coagulation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5158-5162(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS THR-352; GLN-413 AND LYS-445.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 61-466, AND POST-TRANSLATIONAL MODIFICATIONS.
RX   MEDLINE=89088153; PubMed=3264725;
RA   Thim L., Bjoern S., Christensen M., Nicolaisen E.M., Lund-Hansen T.,
RA   Pedersen A.H., Hedner U.;
RT   "Amino acid sequence and posttranslational modifications of human
RT   factor VIIa from plasma and transfected baby hamster kidney cells.";
RL   Biochemistry 27:7785-7793(1988).
RN   [5]
RP   CARBOHYDRATE-LINKAGE SITES SER-112 AND SER-120.
RX   MEDLINE=91250411; PubMed=1904059;
RA   Bjoern S., Foster D.C., Thim L., Wiberg F.C., Christensen M.,
RA   Komiyama Y., Pedersen A.H., Kisiel W.;
RT   "Human plasma and recombinant factor VII. Characterization of O-
RT   glycosylations at serine residues 52 and 60 and effects of site-
RT   directed mutagenesis of serine 52 to alanine.";
RL   J. Biol. Chem. 266:11051-11057(1991).
RN   [6]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=90062160; PubMed=2511201;
RA   Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T., Takao T.,
RA   Shimonishi Y., Iwanaga S.;
RT   "Identification of a disaccharide (Xyl-Glc) and a trisaccharide
RT   (Xyl2-Glc) O-glycosidically linked to a serine residue in the first
RT   epidermal growth factor-like domain of human factors VII and IX and
RT   protein Z and bovine protein Z.";
RL   J. Biol. Chem. 264:20320-20325(1989).
RN   [7]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=91344709; PubMed=2129367;
RA   Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;
RT   "A new trisaccharide sugar chain linked to a serine residue in the
```

```
RT   first EGF-like domain of clotting factors VII and IX and protein Z.";
RL   Adv. Exp. Med. Biol. 281:121-131(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=96175641; PubMed=8598903;
RA   Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,
RA   Konigsberg W.H., Nemreson Y., Kirchhofer D.;
RT   "The crystal structure of the complex of blood coagulation factor
RT   VIIa with soluble tissue factor.";
RL   Nature 380:41-46(1996).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=99126538; PubMed=9925787;
RA   Zhang E., St Charles R., Tulinsky A.;
RT   "Structure of extracellular tissue factor complexed with factor VIIa
RT   inhibited with a BPTI mutant.";
RL   J. Mol. Biol. 285:2089-2104(1999).
RN   [10]
RP   STRUCTURE BY NMR OF 105-145.
RX   MEDLINE=98367502; PubMed=9692950;
RA   Muranyi A., Finn B.E., Gippert G.P., Forsen S., Stenflo J.,
RA   Drakenberg T.;
RT   "Solution structure of the N-terminal EGF-like domain from human
RT   factor VII.";
RL   Biochemistry 37:10605-10615(1998).
RN   [11]
RP   VARIANT GLN-364.
RX   MEDLINE=91300046; PubMed=2070047;
RA   O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,
RA   Meade T.W., Tuddenham E.G.D.;
RT   "Purification and characterization of factor VII 304-Gln: a variant
RT   molecule with reduced activity isolated from a clinically unaffected
RT   male.";
RL   Blood 78:132-140(1991).
RN   [12]
RP   VARIANTS GLN-364 AND PHE-370.
RX   MEDLINE=92340074; PubMed=1634227;
RA   Marchetti G., Patracchini P., Gemmati D., Derosa V., Pinotti M.,
RA   Rodorigo G., Casonato A., Girolami A., Bernardi F.;
RT   "Detection of two missense mutations and characterization of a repeat
RT   polymorphism in the factor VII gene (F7).";
RL   Hum. Genet. 89:497-502(1992).
RN   [13]
RP   VARIANT TYR-238.
RX   MEDLINE=93372811; PubMed=8364544;
RA   Marchetti G., Ferrati M., Patracchini P., Redaelli R., Bernardi F.;
RT   "A missense mutation (178Cys-->Tyr) and two neutral dimorphisms
RT   (115His and 333Ser) in the human coagulation factor VII gene.";
RL   Hum. Mol. Genet. 2:1055-1056(1993).
RN   [14]
RP   VARIANTS.
RX   MEDLINE=94061028; PubMed=8242057;
RA   Takamiya O., Kemball-Cook G., Martin D.M.A., Cooper D.N.,
RA   von Felten A., Meili E., Hahn I., Prangnell D.R., Lumley H.,
RA   Tuddenham E.G.D., McVey J.H.;
RT   "Detection of missense mutations by single-strand conformational
RT   polymorphism (SSCP) analysis in five dysfunctional variants of
RT   coagulation factor VII.";
RL   Hum. Mol. Genet. 2:1355-1359(1993).
RN   [15]
RP   VARIANTS CHARLOTTE GLN-139 AND GLN-212.
RX   MEDLINE=94264305; PubMed=8204879;
RA   Chaing S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,
RA   Roberts H.R., Blajchman M., Monroe D.M., High K.A.;
RT   "Severe factor VII deficiency caused by mutations abolishing the
RT   cleavage site for activation and altering binding to tissue factor.";
RL   Blood 83:3524-3535(1994).
RN   [16]
RP   VARIANT VAL-354.
RX   MEDLINE=95072589; PubMed=7981691;
RA   Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,
RA   Rodeghiero F., Marchetti G.;
RT   "Topologically equivalent mutations causing dysfunctional coagulation
RT   factors VII (294Ala-->Val) and X (334Ser-->Pro).";
RL   Hum. Mol. Genet. 3:1175-1177(1994).
RN   [17]
RP   VARIANT MIE HIS-307.
RX   MEDLINE=95064662; PubMed=7974346;
RA   Ohiwa M., Hayashi T., Wada H., Minamikawa K., Shirakawa S.,
RA   Suzuki K.;
RT   "Factor VII Mie: homozygous asymptomatic type I deficiency caused by
RT   an amino acid substitution of His (CAC) for Arg(247) (CGC) in the
RT   catalytic domain.";
RL   Thromb. Haemost. 71:773-777(1994).
RN   [18]
RP   VARIANT MET-419.
RX   MEDLINE=96247510; PubMed=8652821;
RA   Arbini A.A., Mannucci P.M., Bauer K.A.;
RT   "A Thr359Met mutation in factor VII of a patient with a hereditary
RT   deficiency causes defective secretion of the molecule.";
RL   Blood 87:5085-5094(1996).
RN   [19]
RP   VARIANTS TRP-283; LYS-325; VAL-358; GLN-364; GLU-402 AND GLN-413.
RX   MEDLINE=97001216; PubMed=8844208;
RA   Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,
RA   Lunghi B., Rodeghiero F., Marchetti G.;
RT   "Mutation pattern in clinically asymptomatic coagulation factor VII
RT   deficiency.";
RL   Hum. Mutat. 8:108-115(1996).
RN   [20]
RP   VARIANT VAL-304.
RX   MEDLINE=97037613; PubMed=8883260;
RA   Tamary H., Fromovich Y., Shalmon L., Reich Z., Dym O., Lanir N.,
RA   Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,
RA   Zaizov R., Seligsohn U.;
RT   "Ala244Val is a common, probably ancient mutation causing factor VII
RT   deficiency in Moroccan and Iranian Jews.";
RL   Thromb. Haemost. 76:283-291(1996).
RN   [21]
RP   VARIANT MORIOKA PRO-13.
RX   MEDLINE=98235713; PubMed=9576180;
RA   Ozawa T., Takikawa Y., Niiya K., Ejiri N., Suzuki K., Sato S.,
RA   Sakuragawa N.;
RT   "Factor VII Morioka (FVII L-26P): a homozygous missense mutation in
RT   the signal sequence identified in a patient with factor VII
RT   deficiency.";
RL   Br. J. Haematol. 101:47-49(1998).
RN   [22]
RP   VARIANTS MALTA THR-194 AND VAL-304.
RX   MEDLINE=98112461; PubMed=9452082;
RA   Alshinawi C., Scerri C., Galdies R., Aquilina A., Felice A.E.;
RT   "Two new missense mutations (P134T and A244V) in the coagulation
RT   factor VII gene.";
RL   Hum. Mutat. Suppl. 1:S189-S191(1998).
RN   [23]
RP   VARIANTS ASP-295 AND GLN-413.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).

  Query Match           50.3%;  Score 99;  DB 1;  Length 466;
  Best Local Similarity   48.8%;  Pred. No. 6.5e-09;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
              ||:||  || ||| | |    | |  |: ||:| : |  ||
Db         61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 13
TMG1_HUMAN
```

```
ID   TMG1_HUMAN      STANDARD;      PRT;   218 AA.
AC   O14668;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 1 precursor (Proline-
DE   rich Gla protein 1) (Proline-rich gamma-carboxyglutamic acid protein
DE   1).
GN   PRRG1 OR TMG1 OR PRGP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97404347; PubMed=9256434;
RA   Kulman J.D., Harris J.E., Haldeman B.A., Davie E.W.;
RT   "Primary structure and tissue distribution of two novel proline-rich
RT   gamma-carboxyglutamic acid proteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9058-9062(1997).
CC   -!- TISSUE SPECIFICITY: Highly expressed in the spinal cord.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamate by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF009242; AAB67070.1; -.
DR   HSSP; P00740; 1CFH.
DR   Genew; HGNC:9469; PRRG1.
DR   MIM; 604428; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     20       POTENTIAL.
FT   CHAIN          21    218       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 1.
FT   DOMAIN         21     83       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       84    106       POTENTIAL.
FT   DOMAIN        107    218       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         24     61       GLA-RICH.
FT   DOMAIN        131    135       POLY-PRO.
SQ   SEQUENCE   218 AA;  24947 MW;  26538A61AB0AEB98 CRC64;

  Query Match             46.7%;  Score 92;  DB 1;  Length 218;
  Best Local Similarity   38.6%;  Pred. No. 4.2e-08;
  Matches   17;  Conservative    7;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||  |   :|||:: | |    | |  |:  ||: : |  ||| :
Db         21 ANGFFEEIRQGNIERECKEEFCTFEEAREAFENNEKTKEFWSTY 64


RESULT 14
THRB_RAT
ID   THRB_RAT       STANDARD;      PRT;   617 AA.
AC   P18292;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=90332426; PubMed=2377469;
RA   Dihanich M., Monard D.;
RT   "cDNA sequence of rat prothrombin.";
RL   Nucleic Acids Res. 18:4251-4251(1990).
RN   [2]
RP   SEQUENCE OF 383-617 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
CC       fragment (fragment 1) of the prothrombin, prior to its activation
CC       by factor Xa.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52835; CAA37017.1; -.
DR   EMBL; M81397; AAA42240.1; -.
DR   PIR; S10511; S10511.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR003966; Peptidase_S1A_pr.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
```

```
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   PROPEP      25     43
FT   CHAIN        44    617       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    323       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       324    359       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       360    617       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      360    617       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        323    324       CLEAVAGE (BY FACTOR XA).
FT   SITE        359    360       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    402    402       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    458    458       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    564    564       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD    120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    552    552       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    292       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    287       BY SIMILARITY.
FT   DISULFID    332    478       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    387    403       BY SIMILARITY.
FT   DISULFID    532    546       BY SIMILARITY.
FT   DISULFID    560    590       BY SIMILARITY.
SQ   SEQUENCE   617 AA;  70411 MW;  AD27D1B71445DB1D CRC64;

  Query Match             43.9%;  Score 86.5;  DB 1;  Length 617;
  Best Local Similarity   42.2%;  Pred. No. 1.2e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||:|:| | |:  | :  |    |   ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 15
THRB_MOUSE
ID   THRB_MOUSE     STANDARD;      PRT;   618 AA.
AC   P19221;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2 OR CF2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=91025551; PubMed=2222810;
RA   Friezner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA   Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse prothrombin and
RT   localization of the gene on mouse chromosome 2.";
RL   DNA Cell Biol. 9:487-498(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 384-618 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,
CC       converts fibrinogen to fibrin and activates factors V, VII, VIII,
CC       XIII,.and, in complex with thrombomodulin, protein C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: Prothrombin is activated on the surface of a
CC       phospholipid membrane that binds the amino end of prothrombin and
CC       factors Va and Xa in Ca-dependent interactions; factor Xa removes
CC       the activation peptide and cleaves the remaining part into light
CC       and heavy chains. The activation process starts slowly because
CC       factor V itself has to be activated by the initial, small amounts
CC       of thrombin.
CC   -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal
CC       fragment (fragment 1) of the prothrombin, prior to its activation
CC       by factor Xa.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   -!- SIMILARITY: Contains 2 kringle domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X52308; CAA36548.1; -.
DR   EMBL; BC013662; AAH13662.1; -.
DR   EMBL; M81394; AAA40435.1; -.
DR   PIR; A35827; A35827.
DR   HSSP; P00734; 1B7X.
DR   MEROPS; S01.217; -.
DR   MGD; MGI:88380; F2.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR003966; Peptidase_S1A_pr.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   PROPEP      25     43
FT   CHAIN        44    618       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    324       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       325    360       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       361    618       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      361    618       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        324    325       CLEAVAGE (BY FACTOR XA).
FT   SITE        360    361       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    403    403       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    459    459       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    565    565       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    293       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    288       BY SIMILARITY.
FT   DISULFID    333    479       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    388    404       BY SIMILARITY.
FT   DISULFID    533    547       BY SIMILARITY.
FT   DISULFID    561    591       BY SIMILARITY.
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    413    413       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    553    553       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   618 AA;  70268 MW;  B89F719AAFD601E0 CRC64;

  Query Match             43.9%;  Score 86.5;  DB 1;  Length 618;
  Best Local Similarity   42.2%;  Pred. No. 1.2e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
              ||| ||  ||:|:| | |:  | :  |    |  ||  ||:|:
Db         44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


Search completed: March  1, 2004, 10:03:23
Job time : 10.5 secs
```